

~~PCT/US2005/027239~~

moving from start (N-terminus or 5') to end (C-terminus or 3'), such that for an alignment that extends to p monomers (where $p > x$) there are $p-x+1$ such windows, each window has at least $x \cdot y$ identical aligned monomers, where: x is selected from 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200; y is selected from 0.50, 0.60, 0.70, 0.75, 0.80, 0.85, 0.90, 0.91, 0.92, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98, 0.99; and if $x \cdot y$ is not an integer then it is rounded up to the nearest integer. The preferred pairwise alignment algorithm is the Needleman-Wunsch global alignment algorithm [Needleman & Wunsch (1970) *J. Mol. Biol.* 48, 443-453], using default parameters (e.g., with Gap opening penalty = 10.0, and with Gap extension penalty = 0.5, using the EBLOSUM62 scoring matrix). This algorithm is conveniently implemented in the *needle* tool in the EMBOSS package [Rice *et al.* (2000) *Trends Genet.* 16:276-277].

The nucleic acids and polypeptides of the invention may additionally have further sequences to the N-terminus/5' and/or C-terminus/3' of these sequences (a) to (d).

All of the Gram positive bacterial sequences referenced herein are publicly available through PubMed on GenBank.

Streptococcus pneumoniae Adhesin Island Sequences

As discussed above, a *S. pneumoniae* AI sequence is present in the TIGR4 *S. pneumoniae* genome. Examples of *S. pneumoniae* AI sequences are set forth below.

SrtD (Sp0468) is a sortase. An example of an amino acid sequence of SrtD is set forth in SEQ ID NO: 80.

SEQ ID NO: 80

MSRTKLRALLGYLLMLVACLIPYICFGQMVLSLQSLGQVKGHATFVKSMTTEMYQEQQNHSLAYNQRLASQNRIVDP
FLAEGYEVNYQVSDDPDAVYGYLSIPSLEIMEPVYLGADYHHLGMGLAHVDGTPPLDGTGIRSVIAGHRAEPSH
VFFRHLQDLKVGDALYDNGQEIVEYQMMDEIILPSEWEKLESVSSKNIMTLITCDPIPTFNKRLLVNFERVAV
YQKSDPQTAAVARVAFTKEGQSVSRVATSQWLYRGLVVLAFGLFVLWKLARLLRGK

SrtC (Sp0467) is a sortase. An example of an amino acid sequence of SrtC is set forth in SEQ ID NO: 81.

SEQ ID NO: 81

MSRYYYRIESNEVIKEFDETVSQMDKAELEERWRLAQAFNATLKPSEILDPFTEQEKKGVSSEYANMLKVHERIG
YVEIPAIQEIIPMYVGTSEDILQKGAGLLEGASLPVGGENHTHTVITAHRLPTAELFSQLDKMKKGDIFYLHVLD
QVLAYQVDQIVTVEPNDFEPVLIQHGEDYATLLTCTPYMINSHRLLVRGKRIPYTAPIAERNRAVRERGQFWLWL
LLGAMAVILLLLYRVYRNRRIVKGLEKQLEGRHVKD

SrtB (SP0466) is a sortase. An example of an amino acid sequence of SrtB is set forth in SEQ ID NO: 82.

SEQ ID NO: 82

MAVMAYPLVSRLYRVESNQIADFDKEKATLDEADIDERMKLAQAFNDLNNVSGDPWSEEMKKKGRAEYARM
LEIHERMGHVEIPVIDVDLPVYAGTAEVQLQQGAGHLEGTSPLPIGGENSTHAVITAHGLPTAKMFTDLTKLVGD
KFYVHNIKEVMAYQVDQVKVIEPTNFDDLLIVPGHDYVTLTCTPYMINTHRLLVRGHRIPYVAEVEEEFIAANK
LSHLRYLFFVAVGLIVILLWIIRLRKKKKQPEKALKALKAAARKEVKVEDGQQ

Sp0465 is a hypothetical protein. An example of an amino acid sequence of Sp0465 is set forth in SEQ ID NO: 83.

SEQ ID NO: 83 115 05 / 27 239

MFLPFLSASLYLQTHHFIAFPNRQSYLLRETRKSHFFLIHHPF

RrgC (SP0464) is a cell wall surface anchor family protein. RrgC contains a sortase substrate motif VPXTG (SEQ ID NO: 137), shown in *italics* in SEQ ID NO: 84.

SEQ ID NO: 84

MISRIFFVMALCFSLVWGAHAVQAQEDHTLVLQLENYQEVSQVLSRSGHRLQVWKLDDSYSDRRVQIVRDLHS
WDENKLSSFKKTSFEMTFLENQIEVSHIPNGLYYVRSIIQTDVASYPAEFLFEMTDQTVPEPLVIVAKKTDMTTK
VKLIKVDQDHNRLLEGVGFKLVSARDVSEKEVPLIGEYRYSQGVGRITLYTDKNGEIVFTNLPLGNYRFKEVEP
LAGYAVTTLDTDVQLVDHQLVTITVVNQKLPRGNVDFMKVDGRNTNSIQGAMFKVMKEESGHYTPVLQNGKEVVV
TSGKDGRFRVEGLEYGTYLWELQAPTGYVQLTSPVSFTIGKDRKELVTVVKNNKRPRIDVDPDTGEETLYILML
VAILLFGSGYYLTKKPNN

RrgB (Sp0463) is a cell wall surface anchor protein. RrgB contains a sortase substrate motif IPXTG (SEQ ID NO: 133), shown in *italics* in SEQ ID NO: 85.

SEQ ID NO: 85

MKSINKFLTMLAALLLTASSLFSAAVFAAGTTTTSVTVHKLLATDGDMDKIANELETGNYAGNKVGVLPAKAKE
IAGVMFVWNTNNEIIDENGQTLGVNIDPQTFKLSGAMPATAMKKLTEAEGAKFNTANLPAKYKIYEIHSLSY
VGEDGATLTGSKAVPIEIEPLNDVDAHVYPKNTEAKPKIDKDFKGANPDTPRVDKDTPVNHQVGDVVEYEIV
TKIPALANYATANWSDRMTEGLAFNKGTVKVTVDVALEAGDYALTEVATGFDLKLTDAGLAKVNDQNAEKT VKI
TYSATLNDKAIVEVPESNDVTFNYGNNPDHGNTPKPNKPNENGDLTLTKTWVDATGAPI PAGAEATFDLVNAQTG
KVVQTVTLTDDKNTVTVNGLDKNTYKFVERSISKYSADYQEITTAGETIAVKNWKENPKPLDPTPEPKVVITYGKK
FVKVNDKDNRLAGAEFVIANADNAGQYLARKADKVSQEEKQLVVTTKDALDRAVAAYNALTAQQQTQQEKEKVDK
AQAAAYNAAVIAANNAFEWVADKDNENVKLVSDAQRFEITGLLAGTYYLEETKQPAGYALLTSRQKFEVTATSY
SATGQGIETAGSGKDDATKVVNKKITIPQTGGIGITII FAVAGAAIMGIAVYAYVKNKDEDQLA

RrgA (Sp0462) is a cell wall surface anchor protein. RrgA contains a sortase substrate motif YPXTG (SEQ ID NO: 186), indicated in *italics* in SEQ ID NO: 86.

SEQ ID NO: 86

MLNRETHMKVKRIFQKAVAGLCCISQLTAFSSIVALAETPETSIPAIGKVVIKETGEGGALLGDAVFELKNNTDG
TTVSQRTEAQTGEAIFSNIKPGTYTLTEAQQPVGYKPSKQWTVVEVEKNGRTTVQGEQVENREEALSQYPQTGT
YPDVQTPYQIIKVDGSEKNGQHKALNPNPYERVEPTGLSKRIYQVNNLDDNQYGIELTVSGKTVYEQKDKSVPL
DVVILLDNSNSMSNIRKNARRAERAGEATRSIDKITSDSENVALVTYASTIFDGTFTVEKGVADKNGKRLN
DSLFWNYDQTSFTTNTKDYSYLKLTNDKNDIVELKNKVPTEAEDHDGNRLMYQFGATFTQKALMADEILTQQR
QNSQKVI FHTDGVPTMSYPINFNATFAPSQNQLNAFFSKSPNKDGILLSDFITQATSGEHTIVRGDGSYQM
FTDKTVYEKGAPAAFPVKPEKYSKMAAGYAVIGDPINGGYIWLNWRESILAYPFNSNTAKITNHGDPTRWYNG
NIAPDGYDVFTVGIGINGDPGTDEATATSFMQSISSKPENYTNVTDTTKILEQLNRYFHTIVTEKKSIENTITD
PMGELIDLQLGTDGRFPADYTLTANDGSRLENGQAVGGPQNDGGLLKNKAVLYDTTEKRIRVTVGLYLGTDEKVT
LTYNVRLNDEFVSNKPYDTNGRTTLHPKEVEQNTVRDFPIPKIRDVRKYPEITISKEKKLGDIEFIKVNKNDKKP
LRGAVFSLQKQHPDYPDIYGAIDQNGTYQNVRTGEDGKLTFFKNLSGDKYRLFENSEPAGYKPVQNKPIVAFQIVN
GEVRDVTISIVPQDIPAGYEFTNDKHYITNEPIPKREYPRTGIGMLPFYLIGCMMGGVLLYTRKHP

RlrA (Sp0461) is a transcriptional regulator. An example of an amino acid sequence for RlrA is set forth in SEQ ID NO: 87.

SEQ ID NO: 87

MLNKYIEKRITDKITILNILLDIRSIELDELSTLTSLQSKSLLSILQELQETFEELTFNLDTQQVQLIEHSHQ
TNYFFHQLYNQSTILKILRFFLLQGNQSFNEFTQKEYISIASATGYRVRQKCGLLRSVGLDLVKNQVVGPEYRIRF
LIALQLQHFHIGIEIYDLNDGSMWVTHMIVQSNSQLSHELLEITPDEYVHFSILVALTWKRREFFLEFPESKEFEK
LKNLFMYPILMEHCQTYLEPHANMTFTQEELDYIFLVYCSANSSFSKDKWNQEKKTHTIQLILOHTRGKHLSSKF
KNILGNDISNSLSFLTALTFLTRTFLFGLQNLVPPYNYEYHYGIESDKPLYHISKAIVQEWMTQKIEGVIDQHR
LYLFSLYLTETIFSSLPAPIFIIILNNQADVNLIKSIILRNFTDKVASVTGYNILISPPPSEHLLTEPLIIITTK
EYLPYVKKQYPKGKHHFLTIALDLHVSQQRLIYQITIVDIRKEAFDKRVAMIAKKAHYLL

As discussed above, a *S. pneumoniae* AI sequence is present in the *S. pneumoniae* strain 670 genome. Examples of *S. pneumoniae* AI sequences are set forth below.

~~PCT/US2005/027239~~
 Orf1_670 is a transposase. An example of an amino acid sequence of orf1_670 is set forth in SEQ ID NO: 171.

SEQ ID NO: 171

MEHINHTLLIGIKDKNITLNKAIQHDTHIEVFATLDYHPPCKCHKGKQIKYDFQKPSKIPFIEIGGFPSLIHL
 KKRRFQCKSCRKVTVAETTLVQKNCQISEMVRQKIAQLLLNREALTHIASKLAISTSTVYRKLKQHFQEDYT
 TLPEILSWDEFSYQKGKLAFAQDFNTKKIMTILDNRRQTTIRNHFFKYSKEARKKVKVTVDMMSGYIPLIKKL
 FPNKIVLDRFHIVQHMSRALNQTRINIMKQFDDKSLEYRALKYWKFILKDSRKLSLKPFFYARTFRETLPREC
 LKKIFTLVPKLDYDLYQLLLFHLQEKNTDQFWGLIQDTLPHLNRTFKTTLSTFICYKNYITNAIELPYSNAKL
 EATNKLIKDIKRNAFGFRNFENFKKRIFIALNIKKERTKFFVLSRA

Orf2_670 is a transcriptional regulator. An example of an amino acid sequence of Orf2_670 is set forth in SEQ ID NO: 172.

SEQ ID NO: 172

MLNKYIEKRITDKITILNILLDIRSIELDELSTLTSLQSKSLLSILQELQETFEELTFNLDTOQQVQLIEHHSHQ
 TNYFFHQLYNQSTILKILRFFLLQGNQSFNEFTQKEYISIATGYRVRQKCGLLRSVGLDLVKNQVVGPEYRIRF
 LIALQLFHFGIEIYDLNDGSMWVTHMIVQSNQSLSHELLEITPDEYVHFSILVALTWKRREFPLEFPESKEFEK
 LKNLFMYPILMEHCQTYLEPHANMTFTQEELDYIFLVYCSANSSFSKDKWNQEKKTHTIQLILQHTRGKHLLSKF
 KNILGNDISNSLSFLTALTFLTRTFLFGLQNLVPPYNYEYHYGIESDKPLYHISKAIVQEWMTQKIEGVIDQHR
 LYLFSLYLTETIFSSLPAPIFIFILNNQADVNLIKSIILRNFTDKVASVTGYNILISPPPSEEHLEPLIIITTK
 EYLPYVKKQYPRGKHHFLTIALDLHVSQORLIYQTIIVDIRKEAFDKRVAMIAKKAHYLL

Orf3_670 is a cell wall surface anchor family proten. An example of an amino acid sequence of Orf3_670 is set forth in SEQ ID NO: 173.

SEQ ID NO: 173

MLNRETHMKKVRKIFQKAVAGLCCISQLTAFSSIVALAETPETSIPAIGKVVIKETGEGGALLGDVAFELKNNTDG
 TTVSQRTEAQTGEAIFSNIKPGTYTLTEAQPVGYPKSTQWTVVEVEKNGRTTVQGEQVENREEALSQYPQTGT
 YPDVQTPYQIIKVDGSEKNGQHKALNPNPYERVIPEGLTSKRIYQVNNLDDNQYGIELTVSGKTTVETKEASTPL
 DVVILLDNSNSMSNIRHNHAHRAEKAGEATRALVDKITSNPDNRVALVTYGSTIFDGSEATVEKGVADANGKILN
 DSALWTFDRFTFTAKTNYNSFLNLTSDPTDIQTIKDRIPSDAEELNKDKLMYQFGATFTQKALMTADDILTQOAR
 PNSKKVIFHITDGVPTMSYPINFKYTGTTQSYRTQLNNFKAKTPNSSGILLEDFVTWSADGEHKIVRGDGESYQM
 FTKKPVTDQYGVHQILSITSMEQRAKLVSAGYRFYGTDLYLWYRDSILAYPFNSSTDWITNHDPTTWYNGNMA
 QDGYDVFTVGVGVNGDPTDEATATRFMQSISSPDNYTNVADPSQILQELNRYFYTIVNEKKSIENTGITDPMG
 ELIDFQLGADGRFDPADYTLTANDGSSLVNNVPTGGPQNDGGLLKNAKFYDTTEKRIRVGTGLYLGTGEKVTLTY
 NVRLNDQFVSNKFYDTNGRTTLHPKEVEKNTVRDFFIPKIRDVRKYPEITIPKEKKLGEIEFIKINKNDKKPLRD
 AVFSLQKHDPDYPDIYGAIQNGTYQNVRTGEDGKLTFFKNLSDGKYRLFENSEPAGYKPVQNKPIVAFQIVNGEV
 RDVTSIVPQDIPAGYEFTNDKHYITNEPIPPKREYPRTEGGIGMLPFYLIGCMMMGVLLYTRKHP

Orf4_670 is a cell wall surface anchor family protein. An example of an amino acid sequence of orf4_670 is set forth in SEQ ID NO: 174.

SEQ ID NO: 174

MKSINKFLTMLAALLLTASSLFSAAVFAADNVSTAPDAVTKTLTIHKLLEDDLKTWDTNGPKGYDGTQSSLK
 DLTGVVAEEIPNVYFELQKYNLTDGKEKENLKDDSKWTTVHGGLTTKDGLKIETSTLKGVYRIRREDRTKTTYVGP
 NGQVLTGSKAVPALVTLPLVNNNGTVIDAHVFPKNSYNKPVVDKRIADTLNNDQNGLSIGTKIPYVNTTIPSN
 ATFATSFWSDEMTEGLTYNEDVTITLNNVAMDQADYEVTKGNNGFNKLTEAGLAKINGKDADQKIQITYSATLN
 SLAVADIPESNDITYHYGNHQDHGNTPKPTKPNNGQITVTKTWDSQPAPEGVKATVQLVNAKTGEKVGAPVELSE
 NNWYTWSGLDNSIEYKVEEYNGYSAEYTVESKGLGVKNWKNNDNPAIPNPEEPRVKTYGKKFVKVDQKDTRE
 NAQFVVKKADSNKYIAFKSTAQQAADKAAATAKQKLDAVAAYTNAADKQAAQALVDQAQQEYNVAYKEAKFGY
 VEVAGKDEAMVLTSDTGGQFQISGLAAGTYKLEEIKAPEGFAKIDDSVEFVVGAGSWNQGEFNYLKDVKQNDATKV
 VNKKITIPQTGGIGTIIFAVAGAAIMGIAYYAVKNNKDEDQLA

Orf5_670 is a cell wall surface anchor family protein. An example of an amino acid sequence of orf5_670 is set forth in SEQ ID NO: 175.

SEQ ID NO: 175

MTMQMKQKMISRIFFVMALCFSLVGAHAVQAQEDHTLVLQLENYQEVVSQPLPSRDGHRQLQVWKLDDSYSDDRV
 QIVRDLHSDWENKLSSFKKTSFEMTFLENQIEVSHIPNGLYYVRSIIQTDAVSYPAEFLFEMTDQTVEPLVIVAK

KTDMTTKVKLEIKVDQDHNRLTEGVGTFKLVSVARDGSEKEVPLIGEYRYSSSGQVGRITLYTDKNGEIFVTNLP LGN
 YRFKEVEPLAGYAVTTLDTDVQLVDHQLVTITVVNQKLPRGNVDFMKVDGRTNTSLQGAMFKVMKEESGHYTPVL
 QNGKEVVVTSKDGFRFRVEGLEYGTYLLWELQAPTGYVQLTSPVSFTIGKDKRELVTVVKNNKRPRIDVPDTGE
 ETLYILMLVAILLFGSGYLLTKKPNN

Orf6_670 is a sortase. An example of an amino acid sequence of orf6_670 is set forth in SEQ ID NO: 176.

SEQ ID NO: 176

MLIKMKVTKKKQKRNNLLGVVFFIGMAVMAYPLVSRLYYRVESNQIADFDKEKATLDEADIDERMKLAQAFNDS
 LNNVSGDPWSEEMKKKGRAEYARMLEIHERMGHVEIPVIDVDLPVYAGTAAEVLQAGHLEGTSLPIGGNSTH
 AVITAHTGLPTAKMFTDLTKLVGDKFYVHNIKEVMAYQVDQVKVIEPTNFDDLLIVPGHDYVTLTCTPYMINT
 HRLLVGRHRIPYVAEVEEEFIAANKLSHLRYLFYVAVGLIVILLWIIRLRKKKKQPEKALKALKAAARKEVKVE
 DGQQ

Orf7_670 is a sortase. An example of an amino acid sequence of orf7_670 is set forth in SEQ ID NO: 177.

SEQ ID NO: 177

VSRYYYRIESNEVIKEFDETVSOMDKAELEERWRLAQAFNATLKPSEILDPFTEQEKKGVSSEYANMLKVHERIG
 YVEIPAIQDEIPMYVGTSEEILQKGAGLLEGASLPVGGENTHTTVTAHRLPTAELFSQLDKMKKGDVFFLHVLD
 QVLAYQVDQILTVEPNDFEPVLIQHGEDYATLLTCTPYMINSHRLLVRGKRIPYTAPIAERNRAVRERGGQFWLWL
 LLAALVMILVLSYGVYRHRIRIVKGLEKQLEEHVKG

Orf8_670 is a sortase. An example of an amino acid sequence of orf8_670 is set forth in SEQ ID NO: 178.

SEQ ID NO: 178

MSKAKLQKLLGYLLMLVALVIPVYCFGQMVLSLQVKGHEIFSESVTADSYQEQLQRSILDYNQRLDSQNRIVDP
 FLAEGYEVNYQVSDDDPAVYGYLSIPSLEIMEPVYLGADYHHLAMGLAHVDGTPLPVEGKGIRSVIAGHRAEPSH
 VFFRHLDQLKVGDALYYDNGQEIVEYQMDTEIILPSEWEKLESVSSKNIMTLITCDPIPTFNKRLLVNFERVAV
 YQKSDPQTAAVARVAFTKEGQSVSRVATSQWLYRGLVVLAFGLILFVLWKLARLLRGK

As discussed above, a *S. pneumoniae* AI sequence is present in the 19A Hungary 6 *S. pneumoniae* genome. Examples of *S. pneumoniae* AI sequences from 19A Hungary 6 are set forth below.

ORF2_19AH is a transcriptional regulator. An example of an amino acid sequence of ORF2_19AH is set forth in SEQ ID NO: 187.

SEQ ID NO: 187

MLNKYIEKRITDKITILNILLDIRSIELDELSTLTSLQSKSLLSILQELQETFEELTFNLDLTQQVQLIEHHSHQ
 TNYFFHQLYNQSTILKILRFFLLQGNQSFNEFTQKEYISIASATGYRVRQKCGLLLSVGLDLVKNQVVGPEYRIRF
 LIALQLQFHFGIEIYDLNDGSMWVTHMIVQSNQSLSHELLEITPDEYVHFSILVALTWKRREFPLEFPESKEFEK
 LKNLFMYPILMEHCQTYLEPHANMTFTQEELDYIFLVYCSANSSFSKDKWNQEKKTHTIQLILQHTRGKHLSSKF
 KNILGNDISNSLSFLTALTFLTRTFLGLQNLVPPYNNYEHYHIESDKPLYHISKAIQEWMTQKIEGVIDQHR
 LYLFSLYLTETIFSSLPAPIFIILNNQADVNLIKSIILRNFTDKVASVTGYNILISPPPSEEHLTEPLIIITTK
 EYLPYVKKQYPKGKHHFLTIALDLHVSQQRLLIYQITIVDIRKEAFDKRVAMIAKKAHYLL

ORF3_19AH is a cell wall surface protein. An example of an amino acid sequence of

ORF3_19AH is set forth in SEQ ID NO: 188.

SEQ ID NO: 188

MKKVRKIFQKAVAGLCCISQLTAFSSIVALAETPETSIPAIGKVVIKETGEGGALLGDAVFELKNNTDGTTVSQRT
 EAQTGEAIFSNIKPGTYTLTEAQQPPVGYPSTKQWTVEVEKNGRTTVQGEQVENREEALSQYPTGTYPDVQTP
 YQIKVDGSEKNGQHKALNPNPYERVEPTLSKRIYQVNNLDDNQYGIELTVSGKTTVETKEASTPLDVILLD
 NSNSMSNIRHNHAAHRAEKAGEATRALVDKITSNPDNRVALVTYGSTIFDGSEATVEKGVADANGKILNDSALWTF
 DRTTFTAKTYNYSFLNLTSDPTDIQTIKDRIPSDAEELNKDKLMYQFGATFTQKALMTADDILTKQARPNSKKVI
 FHITDGVPTMSYPINFKYTGTTQSYRTQLNNFKAKTPNSSGILLEDFVTSADGEHKIVRGDGESYQMFTKKPVIT

DYGVHQTLSSTSMEOAKLVSAGYRFGTDLVLYWRDSILAYPFNSSTDWITNHGDPTTWYYNGNMAQDGYDVE
 TVGVGVNGDPGTDEATATRFMQSISSSPDNYTNVADPSQILQELNRYFYTIVNEKKSIENTITDPMGELIDFQL
 GADGRFDPADYTLTANDGSSLVNNVPTGGPQNDGGLLKNKVFYDTTEKRIRVTGLYLGTGEKVTLTYNVRNDQ
 FVSNKFYDTNGRNTLHPKEVEKNTVRDFPIPKIRDVRKYPEITIPKEKKLGEIEFIKINKNDKKPLRDAVFSLQK
 5 QHPDYPDIYGAIDQNGTYQNVRTGEDGKLTfKNLSDGKYRLFENSEPAGYKPVQNKPIVAFQIVNGEVRDVTISV
 PQDIPAGYEFTNDKHYYITNEPIPPKREYPRRTGGIGMLPFYLIGCMMMGGVLLYTRKNP

ORF4_19AH is a cell wall surface protein. An example of an amino acid sequence of
 ORF4_19AH is set forth in SEQ ID NO: 189.

SEQ ID NO: 189

MKSINKFLTMLAALLLTASSLFSAAVFAADNVSTAPDAVTKTLTIHKLLLEDLKTWDTNGPKGYDGTQSSLK
 DLTGVVAEEIPNVYFELQKYNLTGKEKENLKDDSKWTTVHGGLTKDGLKIETSTLKGVIYRIEDRTKTTYVGP
 NGQVLTGSKAVPALVTLPLVNNNGTVIDAHVFPKNSYNKPVVDKRIADTLNNDQNGLSIGTKIPYVNTTIPSN
 ATFATSFWSDEMTEGLTYNEDVTITLNNVAMDQADYEVTGKXNGFNKLKTEAGLAKINGKDADQKIQITYSATLN
 15 SLAVADIPESNDITYHYGNHQDHGNTPKPTKPNNGQITVTKTWDSPAPEGVKATVQLVNAKTGEKVGPVELSE
 NNWTYTWSGLDNSIEYKVEEEYNGYSAEYTVESKGLGVKNWKNNDNNPAPINPEEPRVKTYGKKFVKVDQKDRLE
 NAQFVVKKADSNKYIAFKSTAQQAADEKAAATAKQKLDAVAAYTNAADKQAAQALVDQAQGEYNVAYKEAKFGY
 VEVAGKDEAMVLTSNTDGQFQISGLAAGTYKLEEIKAPEGFAKIDDFEVVVGAGSWNQGEFNYLKDVQKNDATKV
 VNKKITIPQTGGIGTIIFAVAGAAIMGIAVYAYVKNKDEQDLA

ORF5_19AH is a cell wall surface protein. An example of an amino acid sequence of
 ORF5_19AH is set forth in SEQ ID NO: 190.

SEQ ID NO: 190

MTMQKMQMISRIFFVMALCFSLVWGAHAVQAQEDHTLVQLENYQEVVSQLPSPRDGHRQLQVWKLDDSYSDDRV
 25 QIVRDLHSDENKLSFFKTSFEMTFLENQIEVSHIPNGLYYVRSIIQTDVASYPAEFLFEMTDQTVFPLVIVAK
 KTDMTMTTKVLIKVDQDHNRLGEGVGFKLVSVDGSEKEVPLIGEYRYSQVGRITLYTDKNGEIVTNNLPLGN
 YRFKEVEPLAGYAVTTLDTDVQLVDHQLVTITVVNQKLPRGNVDFMKVDGRTNTSLQGAMFKVMKEESGHYTPVL
 QNGKEVVVTSKDGFRFRVEGLEYGTYLLWELQAPTGYVQLTSPVSFTIGKDRKELVTVVKNNKRPRIDVPDTGE
 30 ETLYILMLVAILLFGSGYYLTKKPNN

ORF6_19AH is a putative sortase. An example of an amino acid sequence of ORF6_19AH is
 set forth in SEQ ID NO: 191.

SEQ ID NO: 191

MLIKMVKTKKQRNNLLGVVFFIGMAVMAYPLVSRLYYRVESNQIADFDKEKATLDEADI DERMKLAQAFNDS
 35 LNNVSGDPWSEEMKKKGRAEYARMLEIHERMGHVEIPVIDVDLPVYAGTAEVQLQGAGHLEGTSLPIGNGSTH
 AVITAHTGLPTAKMFTDLTKLVGDKFYVHNIKEVMAYQVDQVKVIEPTNFDDLLIVPGHDYVTLTCTPYMINT
 HRLLVGRHRIPIYVAEVEEEFIAANKLSHLYRYLFYVAVGLIVILLWIIIRLRKKKKQPEKALKALKAAARKEVKVE
 DGQQ

ORF7_19AH is a putative sortase. An example of an amino acid sequence of ORF7_19AH is
 set forth in SEQ ID NO: 192.

SEQ ID NO: 192

MDNSRRSRKKGTKKKHPLILLLIFLVGFAVAIYPLVSRYYYRIESNEVIKEFDETVSQMDKAELEERWRLAQAF
 45 NATLKPSEILDPFTEQEKKGVSSEYANMLKVHERIGYVEIPAIQDEIPMYVGTSEEILQKGAGLLEGASLPVGGE
 NTHTVTAHRLPTAELFSQLDKMKKGDVFYLVHLDQVLAYQVDQILTVEPNDFEPVLIQHGEDYATLLTCTPYM
 INSHRLVGRKRIPTYAPIAERNRAVRERQFWLWLLAALVMILVLSYGVYRHRRIKGLEKQLEEHVKG

ORF8_19AH is a putative sortase. An example of an amino acid sequence of ORF8_19AH is
 set forth in SEQ ID NO: 193.

SEQ ID NO: 193

MSKAKLQKLLGYLLMLVALVIPVYCFGQMVLSQLSGQVKGEHIFSESVTADSYQEQLQRSILDYNQRLDSQNRIVDP
 FLAEGYEVNYQVSDPDVAVGYLSIPSLIMEPVYLGADYHHLAMGLAHVDGTPLPVEGKGIRSVIAGHRAEPSH
 VFFRHLDQLKVGDALYYDNGQEIVEYQMDTEIILPSEWEKLESVSSKNIMTLITCDPIPTFNKRLLVNFERVAV
 55 YQKSDPQTAAVARVAFTKEGQSVSRVATSQWLYRGLVVLAFMGILFVLWKLARLLRGK

As discussed above, a *S. pneumoniae* AI sequence is present in the 6B Finland 12 *S.*

pneumoniae genome. Examples of *S. pneumoniae* AI sequences from 6B Finland 12 are set forth below.

ORF2_6BF is a transcriptional regulator. An example of an amino acid sequence of

ORF2_6BF is set forth in SEQ ID NO: 194.

SEQ ID NO: 194

MLNKYIEKRITDKITILNILLDIRSIELDELSTLTSLQSKSLLSILQELQETFEELTFNLDTQQVQLIEHHSQ
TNYFFHQLYNQSTILKILRFLLQGNQSFNEFTQKEYISIAATGYRVRQKCGLLLRVGLDLVKNQVVGPEYRIRF
LIALLLQFHFGIETIYDLNDGSMDWVTHMIVQSNQSLSHELLEITPDEYVHFSILVALTWKRREFPLEFPESKEFEK
LKNLFMYPILMEHCQTYLEPHANMTFTQEELDYIFLVYCSANSSFSKDKWNQEKKTHTIQLILQHTRGKHLISKF
KNILGNDISNSLSFLTALTFLTRTFLFGLQNLVPYYNYEYHGYIESDKPLYHISKAIQEWMTQEKIEGVIDQHR
LYLFSLYLTETIFSSLPALPIFIILNNQADVNLIKSIILRNFTDKVASVTGYNILISPPPSEEHLEPLIIITTK
EYLPYVKKQYPKGKHHFLTIALDLHVSQQRLIYQTIIVDIRKEAFDKRVAMIAKKAHYLL

ORF3_6BF is a cell wall surface protein. An example of an amino acid sequence of

ORF3_6BF is set forth in SEQ ID NO: 195.

SEQ ID NO: 195

MKKVRKIFQKAVAGLCCISQLTAFSSIVALAETPETSPAIGKVVIKETGEGGALLGDAVFELKNNTDGTTVSQRT
EAQTGEAIFSNIKPGTYTLTEAQQPVGYKPKSTKQWTVVEKNGRTTVQGEQVENREEALSDQYPQTGTYPDVQTP
YQIIKVDGSEKNGQHKALNPNPYERVIPEGTLISKRIYQVNNLDDNQYGIELTVSGKTTVETKEASTPLDVVILL
NSNSMSNIRHNHRAEKAGEATRALVDKITSPDNVALVTYGSTIFDGSEATVEKGVADANGKILNDSALWTF
DRTTFTAKTYNYSFLNLTSDPTDIQTIKDRIPSDAEELNKDKLMYQFGATFTQKALMTADDILTQARPNKVKVI
FHITDGVPTMSYPINFKYTGTTQSYRTQLNNFKAKTPNSSGILLEDFVTWSADGEHKIVRGDGESYQMFTKKPVT
DQYGVHQLISITSMQRAKLVSAGYRFYGTDLVLYWRDSLAYFPNSSTDWITNHGDPPTWYNGNMAQDGYDVF
TVGVGVNGDPGTDEATATRFMQSISSSPDNYTNVADPSQLQELNRYFYTIIVNEKKSIENTITDPMGELIDFQL
GADGRFPADYTLTANDGSSLVNNVPTGGPQNDGGLLNKAKVFYDTTEKRIRVTGLYLGTGEKVTLTYNVRNLDQ
FVSNKFYDTNGRTTLHPKEVEKNTVRDFPIPKIRDVRKYPEITIPKEKKLGEIEFIKINKNDKKPLRDAVFSLOK
QHPDYPDIYGAIDQNGTYQNVRTGEDGKLTFFKNLSDGKYRLFENSEPAGYKPVQNKPIVAFQIVNGEVRDVTISIV
PQDIPAGYEFTNDKHYITNEPIPPKREYPRTGIGMLPFYLGICMMMGVLLYTRKHP

ORF4_6BF is a cell wall surface protein. An example of an amino acid sequence of

ORF4_6BF is set forth in SEQ ID NO: 196.

SEQ ID NO: 196

MKSINKFLTMLAALLLTASSLFSAAATVFAADNVSTAPDAVTKTLTIHKLLSLEDLKTWDTNGPKGYDGTQSSLK
DLTGVAEEIPNVYFELQKYNLTGKEKENLKDDSKWTVHGGTLTKDGLKIETSTLKGVIYRIREDRTKTTYVGP
NGQVLTGSKAVPALVTLPLVNNNGTVIDAHVFPKNSYNKPVVDKRIADTLNNDQNGLSIGTKIPYVNTTIPSN
ATFATSFWSDENTEGLTYNEDVTITLNNVAMDQADYEVTKGNNGFNLKLTAEGLAKINGKDADQKIQITYSATLN
SLAVADIPESNDITYHYGNHQDHGNTPKPTKPNNGQITVTKTWDSQPAPEGVKATVQLVNAKTGEKVGAPVELSE
NNWYTWSGLDNSIEYKVEEYNGYSAEYTVESKGLGVKNWKNNDNPAPINPEEPRVKTYGKKFVKVDQKDRLE
NAQFVVKKADSNKYIAFKSTAQQAADKAAATAKQKLDAAVAAYTNAADKQAAQALVDQAQQEYNAVYKEAKFGY
VEVAGKDEAMVLTSTNDGQFQISGLAAGTYKLEEIKAPEGFAKIDDDVEFVVGAGSWNQGEFNYLKDVQKNDATKV
VNKKITIPQTGGIGTIIFAVAGAAIMGIAVYAYVKNKDEDQLA

ORF5_6BF is a cell wall surface protein. An example of an amino acid sequence of

ORF5_6BF is set forth in SEQ ID NO: 197.

SEQ ID NO: 197

MTMQKMOKMISRIFFVMALCFSLVWGAHAVQAQEDHTLVQLLENYQEVVSQLPSRDGHRQLQVWKLDSDSYSDRV
QIVRDLHSWDENKLSSFKKTSFEMTFLENQIEVSHIPNGLYYVRSIIQTDVSYPAEFLFEMTDQTVPLVIVAK
KTDMTTKVKLIKVDQDHNRLGEGVFKLVSVARDGSEKEVPLIGEYRYSSSGQVGRITLYTDKNGEIVFTNLPLGN
YRFKEVEPLAGYAVTTLDTDVQLVDHQLVTITVVNQKLPRGNVDFMKVDGRINTSLQGAMFKVMKEESGHYTPVL
QNGKEVVVTSKGDGRFRVEGLEYGTYLWELQAPTGYVQLTSPVSTIGKTRKELVTVVKNKRPRIDVPDTGE
ETLYILMLVAILLFGSGYYLTKKPNN

~~PCT/US2005/027239~~
 ORF6_6BF is a putative sortase. An example of an amino acid sequence of ORF6_6BF is set forth in SEQ ID NO: 198.

SEQ ID NO: 198

MLIKMVKTKKQKRNNLLLVVFFIGMAVMAYPLVSRLYYRVESNQIADFDKEKATLDEADI DERMKLAQAFNDS
 LNNVVS GPDWSEEMKKKGRAEYARMLEIHERMGHVEIPVIDVDLPVYAGTAEVVLQGGAGHLEGTSLPIGGNSTH
 AVITAHTGLPTAKMFTDLTKLVGDKFYVHNIKEVMAYQVDQVKVIEPTNFDLLIVPGHDYVTLTCTPYMINT
 HRLLVGRHRI PYVAEVEEEFIAANKLSHLYRYLFYVAVGLIVILLWIIRRLRKKKKQPEKALKALKAAARKEVKVE
 DGQQ

ORF7_6BF is a putative sortase. An example of an amino acid sequence of ORF7_6BF is set forth in SEQ ID NO: 199.

SEQ ID NO: 199

MDNSRRSRKKGTTKKKKHPLILLIFLVGFAVAIYPLVSRYYYRIESNEVIKEFDETVSQMDKAELEERWRLAQAF
 NATLKPSEILDPTFEQEKKGVS EYANMLKVHERIGYVEIPAI DQEI PMYVGTSEEILQKGAGLLEGASLPVGGE
 NTHTVVTAHRLPTAELFSQLDKMKKGDV FYLHVLDQVLAYQVDQILTVEPNDFEPVLIQHGEDYATLLTCTPYM
 INSHLLVRGKRIPYTAPIAERNRAVRERQGFWLWLLLAALVMI LVS YGVYRHRIRIVKGLEKQLEHHVKG

ORF8_6BF is a putative sortase. An example of an amino acid sequence of ORF8_6BF is set forth in SEQ ID NO: 200.

SEQ ID NO: 200

MSKAKLQKLLGYLLMLVALVIPVYCFGQMVLSLQSLGQVKGHEIFSESVTADSYQEQLQSLDYNQRLDSQNRI VDP
 FLAEGYEVNYQVSDDDPAVGYLSIPSLEIMEPVYLGADYHHLAMGLAHVDGTPLPVEGKGIRSVIAGHRAEP SH
 VFFRHL DQLKVG DALYDNGQEIVEYQMMDETIILPSEWEKLESVSSKNIMTLITCDPIPTFNKRLLVNFERVAV
 YQKSDPQTA AVARVAFTKEGQSVSRVATSQWLYRGLVLAFILGILFVLWKLARLLRGK

As discussed above, a *S. pneumoniae* AI sequence is present in the 6B Spain 2 *S. pneumoniae* genome. Examples of *S. pneumoniae* AI sequences from 6B Spain 2 are set forth below.

ORF2_6BSP is a transcriptional regulator. An example of an amino acid sequence of ORF2_6BSP is set forth in SEQ ID NO: 201.

SEQ ID NO: 201

MLNKYIEKRITDKITILNILLDIRSIELDELSTLTSLQSKSILSLQELQETFEELTFNLDTQQVQLIEHHSHQ
 NYYFHLQYNQSTILKILRFFLLQGNQSFNEFTQKEYIS IATGYRVROKCGLLLRVGLDLVKNQVVGPEYRIRF
 LIALLQFHFGIEIYDLNDGSM DWVTHMIVQSNSQLSHELLEITPDEYVHFSILVALTWKRREFPLEFPESKEFEK
 LKNLFMYPILMEHCQTYLEPHANMTFTQEELDYIFLVYCSANSSFSKDKWNQEKKTHTIQLILQHTRGKHL LSKF
 KNILGNDISNLSLFTALTFTLRTFLFGLQNLVPYNYEYHYGIESDKPLYHISKAI VQEWMTQEKIEGVIDQHR
 LYLFSLYLTTETIFSSLP AIPIFIILNNQADVNLIKS IILRNFTDKVASVTGYNILISPPPSEEHLEPLIIITTK
 EYLPYVKKQYPKKGHHFLTIALDLHVSQQRLIYQTIVDIRKEAFDKRVAMI AKKAHYLL

ORF3_6BSP is a cell wall surface protein. An example of an amino acid sequence of

ORF3_6BSP is set forth in SEQ ID NO: 202.

SEQ ID NO: 202

MKKVRKIFQKAVAGLCCISQLTAFSSIVALAETPETS PAIGKVVIKETGEGGALLGDAVFELKNNTDGTTVSQRT
 EAQTGEAIFSNIPGTYTLTEA QPPVGYKPSTKQWTVEVEKNGRTTVQGEQVENREEALSDQYPQTGTYPDVQTP
 YQIIKVDGSEKNGQHAKALNP NPYERVIPEGTLSKRIYQVNNLDDNQYGIELTVSGKTTVETKEASTPLDVILL D
 NSNSMSNIRHNHRAEKA GEATRALVDKITSNPDNRVALVTYGSTIFDGSEATVEKG VADANGKILNDSALWTF
 DRTTFTAKTYNYSFLNLTS DPTDIQTIKDRI PSDAEELNKDKLMYQFGATFTQKALMTADDILTQARPNSKKVI
 FHITDGVPTMSYPINFKYTGTTQSYRTQLN NFKAKTPNSSGILLEDFV TWSADGEHKIVRGDGESYQMFTKKPVT
 DQYGVHQLSITSMEQRAKLVSAGYRFYGTDL YLYWRDSILAYFNSSTDWITNHGDP TTYNGNMAQDG YDVF
 TVGVGVNGDPGTDEATATRFMQSISSSPDNYTNVADPSQILQELNRYFYTIVNEKKS IENGITDPMGELIDFQL
 GADGRFPADYTLTANDGSSLVNNVPTGGPQNDGGLLNKAKV FYDTTEKRIRVTGLYLGTEKVTLTYNVRLNDQ
 FVSNKFYDTNGRTTLHPKEVEKNTVRDFPIPKIRDVRKYPEITIPKEKKLGEIEFIKINKNDKKPLRDAVFSLOK
 QHPDYPDIYGAIDQNGTYQNVRTGEDGKLT FKNLS DGKYLRFENSEPAGYKPQVQNKPIVAFQIVNGEVRDVT SIV
 PQDIPAGYEFTNDKHYITNEPIPPKREYPR TGGIGMLPFYLGICMMMGVLLYTRKHP

~~PCT/US2005/027239~~

ORF4_6BSP is a cell wall surface protein. An example of an amino acid sequence of ORF4_6BSP is set forth in SEQ ID NO: 203.

SEQ ID NO: 203

MKSINKFLTMLAALLLTASSLFSAAATVFAADNVSTAPDAVTKTTLTIHKLLLEDLKTWDTNGPKGYDGTQSSLK
DLTGVAEEI PNVYFELQKYNLTGKEKENLKDDSKWTTVHGGLTKDGLKIETSTLKGVIREDRTKTTYVGP
NGQVLTGSKAVPALVTLPVNNNGTVIDAHVFPKNSYNKPVVDKRIADTLNNDQNGLSIGTKIPYVNTTIPSN
ATFATSFWSDEMTEGLTYNEDVTITLNNVAMDQADYEVTKGNNGFNLKLT EAGLAKINGKDDAQKIQITYSATLN
SLAVADIPESNDITYHYGNHQDHGNTPKPTKPNNGQITVTKTWDSQPAPEGVKATVQLVNAKTGEKVGAPVELSE
NNWYTWSGLDNSIEYKVEEYNGYSAEYTVESKGLGVKNWKNNDNNPAPINPEEPRVKTYGKKFVKVDQKDTRLE
NAQFVVKKADSNKYIAFKSTAQQAADEKAAATAKQKLDAAVAAYTNAADKQAAQALVDQAQQEYNVAYKEAKFGY
VEVAGKDEAMVLTSTNDGQFQISGLAAGTYKLEEKAPGEFAKIDDEFEVVGAGSWNQGEFNYLKDVQKNDATKV
VNKKITIPQTGGIGTII FAVAGAAIMGIAVYAYVKNKDEDQLA

ORF5_6BSP is a cell wall surface protein. An example of an amino acid sequence of

ORF5_6BSP is set forth in SEQ ID NO: 204.

SEQ ID NO: 204

MTMQKMQMISRIFFVMALCFSLVWGAHAVQAQEDHTLVQLQENYQEVVSQLP SRDGHRLQVWKLDDSYSDRRV
QIVRDLHSWDENKLSFKKTSFEMTFLENQIEVSHI PNGLYYVRSIIQTDAVSYPAEFLFEMTDQTVPLVIVAK
KTDTMTTKVKLIKVDQDHNRLLEGVGFKLVS VARDGSEKEVPLIGEYRYSSSGQVGR TLYTDKNGEIFVTNLPLGN
YRFEKEVEPLAGYAVTTLDTDVQLVDHQLVTITVVNQKLP RGNVDFMKVDGRTNTSLQGAMFKVMKEESGHYTPVL
QNGKEVVVTSGKDG RFRVEGLE YGTYYLWELQAPTGYVQLTSPVSFTIGK DTRKELVTVVKNKRPRIDVPDTGE
ETLYILMLVAILLFGSGYYLT KKPNN

ORF6_6BSP is a putative sortase. An example of an amino acid sequence of ORF6_6BSP is set forth in SEQ ID NO: 205.

SEQ ID NO: 205

MLIKMVKTKKQKRNNLLLG VVFFIGMAVMAYPLVSRLYYRVESNQIADFDKEKATLDEADI DERMKLAQAFNDS
LNNVSGDPWSEEMKKKGRAEYARMLEIHERMGHVEI PVIDVDLPVYAGTAEVVLQQGAGHLEGTSLPIGNGSTH
AVITAHTGLPTAKMFTDLTKLVGDKFYVHNIKEVMAYQVDQVKVIEPTNFDDLLIVPGHDYVTLTCTPYMINT
HRLLVGRHRI PYVAEVEEEFIAANKLSHLYRYLFYVAVGLIVILLWII RRLRKKKKQPEKALKALKAAKEVKVE
DGQQ

ORF7_6BSP is a putative sortase. An example of an amino acid sequence of ORF7_6BSP is set forth in SEQ ID NO: 206.

SEQ ID NO: 206

MDNSRRSRKKGTKKKKHPLILLI FLVGFVAIYPLVSRYYYRIESNEVIKEFDETVSQMDKAELEERWRLAQAF
NATLKPSEILDPFTEQEKKGVS EYANMLKVHERIGYVEIPAI DQEI PMYVGTSEEILQKAGLLEGASLPVGGE
NTHTVVTAHRGLPTAELFSQLDKMKKGDFYLVLDQVLAYQVDQILTV EPNDFEPVLIQHGEDYATLLTCTPYM
INSHRLVVRGKRIPYTAPIAERNRAVRERGGFWLWLLLAALVMILVLSYGVYRHRRI VKGLEKQLEEHHVKG

ORF8_6BSP is a putative sortase. An example of an amino acid sequence of ORF8_6BSP is set forth in SEQ ID NO: 207.

SEQ ID NO: 207

MSKAKLQKLLGYLLMLVALVIPVYCFGQMVLSLQGVKGHEIFSESVTADSYQEQLQRSLDYNQRLDSQNRIVDP
FLAEGYEVNYQVSDDPDAVYGYLSIPSL EIMEPVYLGADYHHLAMGLAHVDGTPLPVEGKGIRSVIAGHRAEPSH
VFFRHLDQLKVG DALYYDNGQEI VEYQMMDEIILPSEWEKLESVSSKNIMTLITCDPIPTFNKRLLVNFERVAV
YQKSDPQTAAVARVAFTKEGQSVSRVATSQWLYRGLVVLAF LGILFVLWKLARLLRGK

As discussed above, a *S. pneumoniae* AI sequence is present in the 9V Spain 3 *S. pneumoniae* genome. Examples of *S. pneumoniae* AI sequences from 9V Spain 3 are set forth below.

ORF2_9VSP is a transcriptional regulator. An example of an amino acid sequence of ORF2_9VSP is set forth in SEQ ID NO: 208.

SEQ ID NO: 208

MLNKYIEKRITDKITILNILLDIRSIELDELSTLTSLQSKSLLSILQELQETFEELTFNLDTOQQVQLIEHHSHQ
 TNYFFHQLYNQSTILKILRFFLLQGNQSFNEFTQKEYISIAATGYRVRQKCGLLLRVGLDLVKNQVVGPEYRIRF
 LIALLLQFHFGIEIYDLNDGSMWDVTHMIVQSNQSLSHELLEITPDEYVHFSILVALTWKRREFPLEFPESKEFEK
 5 LKNLFMYPILMEHCQTYLEPHANMTFTQEELDYIFLVYCSANSSFSKDKWNQEKKTHTIQLILQHTRGKHLSSKF
 KNILGNDISNSLSFLTALTFLTRTFLFGLQNLVPYYNYEHYHIESDKPLYHISKAIQVQEWMTQEKIEGVLDQHR
 LYLFSLYLTETIFSSLPAPIFIILNNQADVNLIKSIILRNFTDKVASVTGYNILISPPPSEEHLTEPLIIITTK
 EYLPYVKKQYPKGKHHFLTIALDLHVSQORLIYQITIVDIRKEAFDKRVAMIAKKAHYLL

ORF3_9VSP is a cell wall surface protein. An example of an amino acid sequence of
 ORF3_9VSP is set forth in SEQ ID NO: 209.

SEQ ID NO: 209

MKKVRKIFQKAVAGLCCISQLTAFSSIVALAETPETSIPAIGKVVIKETGEGGALLGDAVFELKNNTNGTTVSQRT
 EAQTGEAIFSNIKPGTYTLTEAQPVPVGYKPKSTKQRTVEVEKNGRTTVQGEQVENREEALSDQYPQTGTYPDVQTP
 15 YQIIKVDGSEKNGQHKALNPYPYERVIPEGTLISKRIYQVNNLDDNQYGIELTVSGKTVYERKDKSVPLDVVILLD
 NSNSMNIERNKNARRAERAGEATRSIDKITSDPENRVALVTYASTIFDGTFTVEKGVADKNGKRLNDSLFWNY
 DQTSFTNTTKDYSYLKLTNDKNDIVELKNKVPTAEADHDGNRLMYQFGATFTQKALMKADEIILTQARQNSQKVI
 FHITDGVPTMSYPINFNHATFAPSQYQNLNAFFSKSPNKDGILLSDFITQATSGEHTIVRGDQSYQMFTDKTVY
 EKGAPAAFPVKPEKYSEMKAAGYAVIGDPIGGYIWLNWRESILAYFNSNTAKITNHGDPTRWYNGNIAPDGY
 20 DVFTVGIGINGDPTDEATATSEMQSISSEKPNYTNVTDTKILEQLNRYFHTIVTEKKSIENGITIDPMGELID
 LQLGTDGRFDPADYTLTANDGSRLENGQAVGGPQNDGGLLKNKAVFYDTEKRIKRVGTGLYLTGKVTLYTNVRL
 NDQFVSNKFYDTNGRTTLHPKEVEKNTVRDFPIPKIRDVRKYPAITIAKEKKLGEIEFIKINKNDKKPLRDAVFS
 LQKQHPDYPDIYGAIDQNGTYQNVRTGEDGKLTFFKNLSDGKYRLFENSEPAGYKPVQNKPIVAFQIVNGEVRDVT
 SIVPQDIPAGYEFTNDKHYITNEPIPPKREYPRGTGGIGMLLFYLGCMMMGGVLLYTRKHP

ORF4_9VSP is a cell wall surface protein. An example of an amino acid sequence of
 ORF4_9VSP is set forth in SEQ ID NO: 210.

SEQ ID NO: 210

MKSINKFLTMLAALLLTASSLSAATVFAAGTTTTSVTVHKKLATDGDMDKIANELETGNYAGNKVGVLPANAKE
 30 IAGVMFVWNTNNEIIDENGQTLGVNIDPQTFKLSGAMPATAMKKLTEAEGAKFNTANLPAAKYKIYEIHSLSY
 VGEDGATLTGSKAVPIEIELPLNDVVDHAVYPKNTEAKPKIDKDFKGANPDTPRVDKDTPVNHQVGDVVEYEIV
 TKIPALANYATANWSDRMTEGLAFNKGTVKVTVDVALEAGDYALTEVATGFDLKLTDAGLAKVNDQNAEKT VKI
 TYSATLNDKAIVEVPESNDVTFNNGNPDHGNTPKPNKPNENGDLTLTKTWVDATGAPIPAGAEATFDLVNAQTG
 KVVQTVTLTDTKNTVTVNGLDKNTYKFFVERSIKGSADYQEITTAGETIAVKNWKDENPKPLDPTEPKVVITYGKK
 35 FVKVNDKDNRLAGAEFVIANADNAGQYLARKADKVSQEEKQLVVTTKDALDRAVAAYNALTAQQQTQQEKEKVDK
 AQAAYNAAVIAANNAFEWVADKDNENVKLVSDAQGRFEITGLLAGTYYLEETKQPAGYALLTSRQKFEVTATSY
 SATGQGIETAGSGKDDATKVVNKKITIPQTGGIGITIFAVAGAVIMGIAVYAYVKNKDEQDLA

ORF5_9VSP is a cell wall surface protein. An example of an amino acid sequence of
 ORF5_9VSP is set forth in SEQ ID NO: 211.

SEQ ID NO: 211

MTMQMKQMKQMKQMKQMKMISRIFFVMALCFSLVWGAHAVQAQEDHTLVQLQENYQEVVSQPSRDGHRLLQVW
 KLDDSYSDNRVQIVRDLHSDWENKLSSEKTSFEMTFLENQIEVSHIPNGLYYVRSIIQTDAVSYPAEFLFEMT
 45 DQTVPEPLVIVAKKADTVTTKVKLIKVDQDHNRLGEGVGFKLVSVDARGSEKEVPLIGEYRYSSESGQVGRITLYTDKN
 GEIVVTNLPLGTYRFEVEPLAGYTVTMDTDVQLVDHQLVTITVNVQKLPRGNVDFMKVDGRNTSLQGAMFKV
 MKEENGHYTPVLQNGKEVVVASGKDGFRFRVEGLEYGTYLWELQAPTGYVQLTSPVSFTIGKDRKELVTVVKNN
 KRPRIDVPDTGEETLYILMLVAILLFGSGYYLTKKTNN

ORF6_9VSP is a putative sortase. An example of an amino acid sequence of ORF6_9VSP is
 set forth in SEQ ID NO: 212.

SEQ ID NO: 212

MLIKMAKTKKQKRNNLLLGVVFFIGIAVMAYPLVSRLYYRVESNQIADFDKEKATLDEADIDERMKLAQAFNDS
 LNNVVSQDPWSEEMKKKGRAEYARMLEIHERMGHVEIPAIDVDLPVYAGTAEEVLQQAGHLEGTSLPIGGNSTH
 55 AVITAHTGLPTAKMFTDLTKLVGDKFYVHNIKEVMAYQVDQVKVIEPTNFDDLLIVPGHDYVTLTCTPYMINT
 HRLVRGHRIPYVAEVEEFIAANKLSHLYRYLFYVAVGLIVILLWIIRRLRKKKRQSERALKALKEATKEVKVE
 DE

PCT/US05/27239

ORF7_9VSP is a putative sortase. An example of an amino acid sequence of ORF7_9VSP is set forth in SEQ ID NO: 213.

SEQ ID NO: 213

MSKSRYSRKKSVKKKNPFILLIIFLVGLAVAMYPLVSRYYYRIESNEVIKEFDETVSQMDKAELEERWRLAQAF
NATLKPSEILDPFTEQEKKGVSSEYANMLKVHERIGYVEIPAIHQEIPMYVGTSEEILQKGAGLLEGASLPVGGG
NTHTVVTAHRLPTAELEFSQDKMKKGDIFYLHVLDQVLAYQVDQIVTVEPNDFEPVLIQHGEDYATLLTCTPYM
INSHRLLVGRKRIPTYAPIAERNRAVRERGGFWLWLLLGAMAVILLLLLYRVYRNRIRVKGLEKQLEGRHVKD

ORF8_9VSP is a putative sortase. An example of an amino acid sequence of ORF8_9VSP is set forth in SEQ ID NO: 214.

SEQ ID NO: 214

MSRTKLRLALLGYLLMLVACLIPYICFGQMVLSLQGVKGHATFVKSMTTEMYQEQQNHSLAYNQRLASQNRIVDP
FLAEGYEVNYQVSDPDPAVYGYLSIPSLEIMEPVYLGADYHHLGMGLAHVDGTPLPLDGTGIRSVIAGHRAEPSH
VFFRHLQDKLVGDALYDNGQEIYEQMMDTEIILPSEWEKLESVSSKNIMTLITCDPIPTFNKRLLVNFERVAV
YQKSDPQTAARVAFAFTKEGQSVSRVATSQWLYRGLVVLAFGLILFVLWKLARLLRGK

As discussed above, a *S. pneumoniae* AI sequence is present in the 14 CSR 10 *S. pneumoniae* genome. Examples of *S. pneumoniae* AI sequences from 14 CSR 10 are set forth below.

ORF2_14CSR is a transcriptional regulator. An example of an amino acid sequence of ORF2_14CSR is set forth in SEQ ID NO: 215.

SEQ ID NO: 215

MLNKYIEKRITDKITILNILLDIRSIELDELSTLTSLQSKSLLSILQELQETFEELTFNLDLTQQVQLIEHHSHQ
TNYFFHQLYNQSTILKILRFFLLQGNQSFNEFTQKEYISIASATGYRVRQKCGLLRSVGLDLVKNQVVGPEYRIRF
LIALLOFHFGIEIYDLNDGSDMWVTHMIVQNSQLSHELLEITPDEYVHFSILVALTWKRREFPLEFPESKEFEK
LKNLFMYPIILMEHCQTYLEPHANMTFTQEELDYIFLVYCSANSSFSKDKWNQEKKTHTIQLILQHTRGKHLISKF
KNILGNDISNSLSFLTALTFLTRTFLFGLQNLVPPYNYEYHYGIESDKPLYHISKAVQEWMTQKIEGVIDQHR
LYLFSLYLTETIFSSLPAPIFIILNNQADVNLKSIILRNFTDKVASVTGYNILISPPPSEHLTEPLIIITTK
EYLPYVKKQYPKGKHHFLTIALDLHVSQQRLLIYQTTIVDIRKEAFDKRVAMIAKKAHYLL

ORF3_14CSR is a cell wall surface protein. An example of an amino acid sequence of ORF3_14CSR is set forth in SEQ ID NO: 216.

SEQ ID NO: 216

MKKVRKIFQKAVAGLCCISQLTAFSSIVALAETPETSIPAIGKVVIKETGEGGALLGDAVFELKNNTDGTTVSQRT
EAQTGEAIFSNIKPGTYTLTEAQPVPVGYKPKSTQWTVVEVEKNGRTTVQGEQVENREEALSDQYPQTGTYPDVQTP
YQIIKVDGSEKNGQHKALNPNPYERVIPEGTLISKRIYQVNNLDDNQYGIELTVSGKTTVETKEASTPLDVILLD
NSNSMSNIRHNHAAKAGEATRALVDKITSNPDNRVALVTYGSTIFDGSEATVEKGVADANGKILNDSALWTF
DRTTFTAKTYNYSFLNLTSDPTDIQTIKDRIPSDAEELNKDKLMYQFGATFTQKALMTADDILTQARPNSKKVI
FHTDGVPTMSYPINFKYTGTTQSYRTQLNNFKAQTPNSSGILLEDFVTSADGEHKIVRGDGESYQMFTKKPV
DQYGVHQILSITSMEQRAKLVSAGYRFYGTDLVLYWRDSILAYPFNSSTDWITNHGDPPTWYNGNMAQDGYDVF
TVGVGVNGDPGTDEATATRFMQSISSSPDNYTNVADPSQILQELNRYFYTIVNEKKSIEGTITDPMGELIDFQL
GADGRFDPADYTLTANDGSSLVNNVPTGGPQNDGGLLKNKAVFYDTTEKRIRVTGLYLGTGEKVTLTYNVRLNDQ
FVSNKFYDTNGRTTLHPKEVEKNTVRDFPIPKIRDVRKYPEITIPKEKKLGEIEFIKINKNDKKPLRDAVFSLQK
QHPDYPDIYGAIDQNGTYQNVRTGEDGKLTFFKNLSGKYRLFENSEPAGYKPVQNKPIVAFQIVNGEVRDVTISV
PQDIPAGYEFTNDKHYITNEPIPPKREYPRGTGGIGMLPFYLLIGCMMMGVLLYTRKHP

ORF4_14CSR is a cell wall surface protein. An example of an amino acid sequence of ORF4_14CSR is set forth in SEQ ID NO: 217.

SEQ ID NO: 217

MKSINKFLTMLAALLLTASSLFSAAATVFAADNVSTAPDAVTKTLTIHKLLESDDLKTWDTNGPKGYDGTQSSLK
DLTGVAEEIPNVYFELQKYNLTGKEKENLKDDSKWTTVHGGTLTKDGLKIETSTLKGVIYRIREDRTKTTYVGP
NGQVLGTGSKAVPALVTLPLVNNNGTVIDAHVFPKNSYNKPVVDKRIADTLNNDQNGLSIGTKIPYVNTTIPSN
ATFATSFWSDEMTEGLTYNEDVTITLNNVAMDQADYEVTKGNNGFNLKLTAEAGLAKINGKDADQKIQITYSATLN
SLAVADIPESNDITYHYGNHQDHGNTPKPTKPNNGQITVTKTWDSQPAPEGVKATVQLVNAKTGEKVGAPVELSE

NWLYTWSGLDNSLEYKVEEHNQYSAEYTVESKGLGVKNWKDNNPAPINPEEPRVKTYGKKFVKVDQKDTRLE
 NAQFVVKKADSNKYIAFKSTAQQAADEKAAATAKQKLDAAVAAYTNAADKQAAQALVDQAQQEYNVAYKEAKFGY
 VEVAGKDEAMVLTSNTDGGQFQISGLAAGTYKLEEIKAPEGFAKIDDEVEFVVGAGSWNQGEFNYLKDVQKNDATKV
 VNKKITIPQTGGIGITIIFAVAGAAIMGIAVYAYVKNKDEDDQLA

ORF5_14CSR is a cell wall surface protein. An example of an amino acid sequence of
 ORF5_14CSR is set forth in SEQ ID NO: 218.

SEQ ID NO: 218

MTMQKMQMISRIFFVMALCFSLVWGAHAVQAQEDHTLVQLQENYQEVVSQLPSRDGHRQLQVWKLDDSYSDRV
 QIVRDLHSWDENKLSSFKKTSFEMTFLENQIEVSHIPNGLYYVRSIIQTDAVSYPAEFLFEMTDQTVFPLVIVAK
 KTDMTTKVKLIKVDQDHNRLLEGVGFKLVSVDGSEKEVPLIGEYRYSSSGQVGRITLYTDKNGEIVFTNLPLGN
 YRFKEVEPLAGYAVTTLDTDVQLVDHQLVTITVVNQKLPGRNVDFMKVDGRTNTSLQGAMFKVMKEESGHYTPVL
 QNGKEVVVTSKDGRRFRVGELEYGTYYLWELQAPTGYVQLTSPVSFTIGKDKTRKELTVVKNKRPRIDVPDTGE
 ETLYILMLVAILLFGSGGYLTKKPNN

ORF6_14CSR is a putative sortase. An example of an amino acid sequence of ORF6_14CSR
 is set forth in SEQ ID NO: 219.

SEQ ID NO: 219

MLIKMKVTKKKQRNNLLLGVVFFIGMAVMAYPLVSRLYRVESNQIADFDKEKATLDEADIDERMKLAQAFNDS
 LNNVSGDPWSEEMKKKGRAEYARMLEIHERMGHVEIPVIDVDLPVYAGTAEVLQQGAGHLEGTSLPIGGNSTH
 AVITAHTGLPTAKMFTDLTKLVGDKFYVHNIKEVMAYQVDQVKVIEPTNFDDLLIVPGHDYVTLTCTPYMINT
 HRLLVGRHRIPIYAEVEEEFIAANKLSHLYRYLFYVAVGLIVILLWIIRRLRKKKKQPEKALKALKARKEVKVE
 DGQQ

ORF7_14CSR is a putative sortase. An example of an amino acid sequence of ORF7_14CSR
 is set forth in SEQ ID NO: 220.

SEQ ID NO: 220

MDNSRRSRKKGTKKKKHPLILLIFLVGFAVAIYPLVSRYYYRIESNEVIKEFDETVSQMDKAELEERWRLAQAF
 NATLKPSEILDPFTEQEKKGVSSEYANMLKVHERIGYVEIPAIQDQEIIPMYVGTSEEILQKGAGLLEGASLPVGGE
 NTHTVVTAHRGLPTAELEFSQLDKMKGDVFYLHVLDQVLAQVDQILTVEPNDFEPVLIQHGEDYATLLTCTPYM
 INSHRLLVRGKRIPYTAPIAERNRAVRERQGFWLWLLAALVMILVLVSYGVYRHRRIKGLQKLEEHVKG

ORF8_14CSR is a putative sortase. An example of an amino acid sequence of ORF8_14CSR
 is set forth in SEQ ID NO: 221.

SEQ ID NO: 221

MSKAKLQKLLGYLLMLVALVIPVYCFGQMVQLQSLGQVKGHEIFSESVTADSYQEQLQRLDYNQRLDSQNRIVDP
 FLAEGYEVNYQVSDDDPDVYGYLSIPSLIMEPVYLGADYHHLAMGLAHVDGTPLPVEGKGIRSVIAGHRAEPSH
 VFFRHLQKLVGDALYYDNGQEIVEYQMMDTETIILPSEWEKLESVSSKNIMTLITCDPIPTFNKRLLVNFERVAV
 YQKSDPQTAARVAVFTKEGQSVSRVATSQWLYRGLVVLAFGLILFVLWKLARLLRGK

As discussed above, a *S. pneumoniae* AI sequence is present in the 19F Taiwan 14 *S.*
pneumoniae genome. Examples of *S. pneumoniae* AI sequences from 19F Taiwan 14 are set forth
 below.

ORF2_19FTW is a transcriptional regulator. An example of an amino acid sequence of
 ORF2_19FTW is set forth in SEQ ID NO: 222.

SEQ ID NO: 222

MLNKYIEKRITDKITILNILLDIRSIELDELSTLTSLQSKSLLSILQELQETFEELTFNLDQVQLIEHSHQ
 TNYFHLQYNQSTILKILRFFLLQGNQSFNEFTQKEYISIAATGYRVRQKCGLLRSVGLDLVKNQVVGPEYRIRF
 LIALQHFHFGIEIYDLNDGSMWVTHMIVQSNQSLSHELLEITPDEYVHFSILVALTWKRREFFLEFPESKEFEK
 LKNLFMPILMEHCQTYLEPHANMTFTQEELDYIFLVYCSANSSFSKDKWNQEKKTHTIQLILQHTRGKHLSSKF
 KNILGNDISNSLSFLTALTFLTRTFLGLQNLVPYYNYEYHYGIESDKPLYHISKAIVQEWMTQKIEGVIDQHR
 LYLFSLYLTETIFSSLPAPIFIIILNNQADVNLKSIILRNFTDKVASVTGYNILISPPPSEEHLTEPLIITTK
 EYLPYVKKQYPKGGKHHFLTIALDLHVSQQRLLIYQITIVDIRKEAFDKRVAMIAKKAHYLL

PCT/US05/27239

ORF3_19FTW is a cell wall surface protein. An example of an amino acid sequence of ORF3_19FTW is set forth in SEQ ID NO: 223.

SEQ ID NO: 223

5 MKKVRKIFQKAVAGLCCISQLTAFSSIVALAETPETSIPAIGKVVIKETGEGGALLGDAVFELKNNTDGTTVSQRT
EAQTGEAIFSNIKPGTYTLTEAQQPPVGYPSTKQWTVVEVEKNGRTTVQGEQVENREEALSQYPQTGYPDVQTP
YQIIKVDGSEKNGQHKALNPNPYERVIPEGTLISKRIYQVNNLDDNQYGIELTVSGKTVYERKDKSVPLDVVILLD
10 NSNSMSNIRNKNARRAERAGEATRSIDKITSDPENRVALVTYASTIFDGTEFTVEKGVADKNGKRLNDSLFWNY
DQTSFTTNTKDYSYLKLTNDKNDIVELKNKVPTEAEDHDGNRLMYQFGATFTQKALMKADEILTQARQNSQKVI
FHITDGVPMTSPINFNHATFAPSYQNQLNAFFSKSPNKDGILLSDFITQATSGEHTIVRGDQGSYQMFTDKTVY
EKGAPAAFPVKPEKYSEMKAAGYAVIGDPINGGYIWLNWRESILAYPFNSNTAKITNHGAPTRWYNGNIAPDGY
DVFTVGIGINGDPGTDEATATSMQSISSKPENYTNVTDTTKILEQLNRYFHTIVTEKKSIENTITDPMGELID
LQLGTDGRFDPADYTLTANDGSRLENGQAVGGPQNDGGLLKNAKFYDTEKRIRVTGLYLGTGEKVTLTYNVRL
15 NDQFVSNKFYDNGRRTLHPKEVEKNTVRDFPIPKIRDVRKYPAITIAKEKKLGEIEFIKINKNDKKPLRDAVFS
LQKQHPDYPDIYGAIDQNGTYQNVRTGEDGKLTFFKNLSDGKYRLFENSEPAGYKPVQNKPIVAFQIVNGEVRDVT
SIVPQDIPAGYEFTNDKHYITNEPIPPKREYPRTEGGIGMLPFYILGCMMSGVLLYTRKHP

ORF4_19FTW is a cell wall surface protein. An example of an amino acid sequence of ORF4_19FTW is set forth in SEQ ID NO: 224.

SEQ ID NO: 224

20 MKSINKFLTMLAALLLTASSLFSAAATVFAAGTTTTSVTVHKLATDGDMDKIANELETGNYAGNKVGVLPANAKE
IAGVMFVWNTNTNNEIIDENGQTLGVNIDPQTFKLSGAMPATAMKKLTEAEGAKFNTANLPAKYKIYEIHSLSY
VGEDGATLTGSKAVPIEIELPLNDVVDHVPKNTAKPKIDKDFKGKANPDTPRVKDTFVNHQVGDVVEYEV
TKIPALANYATANWSDRMTEGLAFNKGTVKVTVDVDALEAGDYALTEVATGFDLKLTDAGLAKVNDQNAEKT VKI
25 TYSATLNDKAIVEVPESNDVTFNYGNPNPDHGNTPKPNKPNENGDLTLTKTWVDATGAPI PAGAEATFDLVNAQTG
KVVQTVTLTTDKNTVTVNGLDKNTYKFFVERSIKGYSAQYQELTTAGEIAVKNWKDENPKPLDPTEPKVVTYGKK
FVKVNDKDNRLAGAEFVIANADNAGQYLARKADKVSQEEKQLVVTTKDALDRAVAAYNALTAQQQTQQEKEKVDK
AQAAYNAAVIAANNAFEWVADKDNENVVKLVSDAQGRFEITGLLAGTYYLEETKQPAYALLTSRQKFEVTATSY
30 SATGQGIETAGSGKDDATKVVNKKITIPQTGGIGTIIFAVAGAVIMGIAYVAYVKNKDEDQLA

ORF5_19FTW is a cell wall surface protein. An example of an amino acid sequence of ORF5_19FTW is set forth in SEQ ID NO: 225.

SEQ ID NO: 225

35 MTMQKMQMISRIFFVMALCFSLVWGAHAVQAQEDHTLVLOLENYQEVVSQLP SRDGHRLQVWKLDDSYSDNRV
QIVRDLHSWDEKLSSEFKKTSFEMTFLNQIEVSHIPNGLYYVRSIIQTDVSYPAEFLFEMTDQTVPLVIVAK
KADTVTTKVKLIKVDQDHNRLLEGVGFKLVSVDGSEKEVPLIGEYRYSSTGQVGRITLYTDKNGEIVVTNLPLGT
YRFKEVEPLAGYTVTMDTDVQLVDHQLVTITVVNQKLPNGVDFMKVDGRTNTSLQGAMFKVMKEENGHYTPVL
QNGKEVVVASGKDGREFRVEGLEYGTYLWELQAPTGYVQLTSPVSFTIGKDRKELVTVVKNNKRPRIDVPDTGE
40 ETLYILMLVAILLFGSGYYLTKKTN

ORF6_19FTW is a putative sortase. An example of an amino acid sequence of ORF6_19FTW is set forth in SEQ ID NO: 226.

SEQ ID NO: 226

45 MLIKMAKTKKQKRNNLLGVVFFIGMAVMAYPLVSRLYRVESNQIADFDKEKATLDEADIDERMKLAQAFNDS
LNNVSGDPWSEEMKKKGRAEYARMLEIHERMGHVEIPADVDLPVYAGTAEVLQQGAGHLECTSLPIGGNSTH
AVITAHTGLPTAKMFTDLTKLVGDKFYVHNIKEVMAYQVDQVKVIEPTNFDDLIVPGHDYVTLTCTPYMINT
HRLVLRGHRIPYVAEVEEEFIAANKLSHLRYLFYVAVGLIVILLWIIRRLRKKRQSERALKALKEATKEVKVE
DE

ORF7_19FTW is a putative sortase. An example of an amino acid sequence of ORF7_19FTW is set forth in SEQ ID NO: 227.

SEQ ID NO: 227

MSKSRYSRKKS VKKKKNPFILLIFLVGLAVAMYPLVSRYYYRIESNEVIKEFDETVSQMDKAELEERWRLAQAF
NATLKPSEILDPFITDQEKKGQVSEYANMLKVHERIGYVEIPAIEQEIIPMYVGTSEDILQKGAGLLEGASLPVGG

NPFLVITAHNGLETALLETSSQIDKMKKGGDTFFLHVLDQVLAQVDQIVTVEPNDFEPVLIQHGQDYATLLTCTPYM
INSHRLLVGRKRIPTYTAPIAERNRAVRERGGQFWLWLLLGAMAVILLLLYRVYRNRRIRVKGLEKQLEGRHVKD

ORF8_19FTW is a putative sortase. An example of an amino acid sequence of

5 ORF8_19FTW is set forth in SEQ ID NO: 228.

SEQ ID NO: 228

MSR TKLRALLGYLLMLVACLIPIYCFGQMVLSLQSLGQVKGHATFVKSM TTEMYQEQQNHSLAYNQRLASQNRIVDP
FLAEGYEVNYQVSDDDPAVYGYLSIPSLEIMEPVYLGADYHHLGMGLAHVDGTPPLDGTGIRSVIAGHRAEPSH
VFFRHLDQLKVGDALYDNGQEIVEYQMM DTEIILPSEWEKLESVSSKNIMTLITCDPIPTFNKRLLVNFERVAV
10 YQKSDPQTAARVARVAFTKEGQSVSRVATSQWLYRGLVVLAFILGILFVLWKLARLLRGK

As discussed above, a *S. pneumoniae* AI sequence is present in the 23F Taiwan 15 *S.*

pneumoniae genome. Examples of *S. pneumoniae* AI sequences from 23F Taiwan 15 are set forth below.

15 ORF2_23FTW is a transcriptional regulator. An example of an amino acid sequence of

ORF2_23FTW is set forth in SEQ ID NO: 229.

SEQ ID NO: 229

MLNKYIEKRITDKITILNILLDIRSIELDELSTLTSLQSKSLLSILQELQETFEELTFNLD TQQVQLIEHHSHQ
TNYFFHQLYNQSTILKILRFFLLQGNQSFNEFTQKEYIS IATGYRVRQKCGLLLR SVGLDLVKNQVVGPEYRIRF
20 LIALQLFHFGIEIYDLNDGSM DWVTHMIVQSNQSLSHELLEITPDEYVHFSILVALTWKRREFPLEFPESKEFEK
LKNLFMYPILMEHCQTYLEPHANMTFTQEELDYIFLVYCSANSSFSKDKWNQEKKTHTIQLILQHTRGKHLLSKF
KNILGNDISNSLSFLTALTFLTRTFLFGLQNLVPYNNYEHYGIESDKPLYHISKAI VQEWMTQEKIEGVIDQHR
LYLFSLYLTETIFSSLPAPIFIILNNQADVNLIKSII LRNFTDKVASVTGYNILISPPPSEEHLEPLIIITTK
EYLPYVKKQYPKGKHHFLTIALDLHVSQQRLIYQTI VDIRKEAFDKRVAMI AKKAHYLL
25

ORF3_23FTW is a cell wall surface protein. An example of an amino acid sequence of

ORF3_23FTW is set forth in SEQ ID NO: 230.

SEQ ID NO: 230

MKKVRKIFQKAVAGLCCISQLTAFSSIVALAETPETS PAIGKVVIKETGEGGALLGDAVFELKNNTDGTTVS QRT
30 EAQTGEAIFSNIKPGTYTLTEA QPPVGYKPKSTQWTVVEKNGRTTVQGEQVENREEALS DQYPQTGTYPDVQTP
YQIIKVDGSEKNGQHKALNPNPYERVIPEG TSLSKRIYQVNNLDDNQYGIELTVSGKTVYE QDKSVPLD VVILLD
NSNSMSNIRNKNARRAERAGEATRSLIDKITS DPENRVALVTYASTIFDGTEFTVEKGVADKNGKRLNDSLF WNY
DQTSFTTNTKDYSLKLTNDKNDIVELKNKVPTEAEDHDGNRLMYQFGATFTQKALMKADEILTQQARONS QKVI
35 FHITDGVPTMSYPINFNHATFAPSYQNQLNAFFSKSPNKDGILLSDFITQATS GEHTIVRGDGQSYQMFTDKTVY
EKGAPAAFPVKPEKYSEMKAAGYAVIGDPINGGYIWLNWRESILAYPFNSNTAKITNHGDPTRWYNGNIAPDGY
DVFTVGIGINGDPGTDEATATSFMQSISSK PENYTNVTDTKILEQLNRYFHTIVTEKKS IENGITIDPMGELID
LQLGTDGRFDPADYTLTANDGSRLENGQAVGGP QNDGGLLKNAKVLYDTTEKRIRV TGLYLGTD EKVTLTYNVRL
NDEFVSNKFYDTNGRTTLHPKEVEQNTVRDFPI PKIRDVRKYPEITISKEKKLGDI EFKVNKNDDKKPLRDAVFS
40 LQKQHPDYPDIYGAIDQNGTYQNVRTGEDG KLTFFKNLSDGKYRLFENSEPAGYKPVQNKPIVAFQIVNGEVRDVT
SIVPQDIPAGYEFNDKH YITNEPIPPKREYPR TGGIGMLPFYLGMMMGVLLYTRKHP

ORF4_23FTW is a cell wall surface protein. An example of an amino acid sequence of

ORF4_23FTW is set forth in SEQ ID NO: 231.

SEQ ID NO: 231

MKSINKFLTILAAALLLVSSLSAATVFAAEQKTKTLTVHKLLMTDQELDAWN SDAITTAGYDGSQNF EQFKQLQ
45 GVPQGVTEISGVAFELQSYTGPGKEQENLTNDAVWTA VNKGVT TETGVKFDTEVLQGT YRLVEVRKESTYVGPN
GKVL TGMAVPALITLPLVNQNGVVENAHVYPKNS EDKPTATKTFDTAAGFVDPGEKGLAIGTKVPYIVTTTIPK
NSTLATAFWSDEMTEGLDYNGDVVN YNGQPLDNSHYTLEAGHNGFILKLNKGLEAINGKDAEATITLKYTATL
NALAVADVPEANDVTFFHYGNNPGHGNTPKPNKPKNGELTITKTWADAKDAPIAGVEVTFDLVNAQTGEVVKVPGH
50 ETGIVLNQTNNTFTATGLDNNT EYKFVERTIKGYSADYQTITETGKIAVKNWKDENPEPINPEEPRVKTYGKKF
VKVDQKDERLKEAQFVVKNEQGKYLALKSAAQQAVNEKAAAEAKQALDAAIAAYTNAADKNAAQAVVDAAQKTYN
DNYRAARFGYVEVERKEDALVLT SNTDGQFQISGLAAGSYTLEETKAPEGFAKLGDVKFEVGAGSWNQGDFNYLK
DVQKNDATKVVNKKITIPQTGGIGITII FAVAGAVIMGIAYVAYVKNKDEDDQLA

ORF5_23FTW is a cell wall surface protein. An example of an amino acid sequence of ORF5_23FTW is set forth in SEQ ID NO: 232.

SEQ ID NO: 232

MTMQMKQKMISRIFFVMALCFSLVWGAHAVQAQEDHTLVQLQENYQEVVSQLPSRDGHRQLQVWKLDDSYSDNRV
 QIVRDLHSDENKLSSFKKTSFEMTFLNQIEVSHIPNGLYYVRSIIQTDAVSYPAEFLFEMTDQTVPLVIVAK
 KADTVTTKVKLKVDQDHNRLGEGVGFKLVSVDGSEKEVPLIGEYRYSSSGQVGRITLYTDKNGEIVVTNLPLGT
 YRFKEVEPLAGYTVTTMDTDVQLVDHQLVTITVVNQKLPRGNVDFMKVDGRTNTSLQGAMFKVMKEENGHYTPVL
 QNGKEVVVASGKDGFRVEGLEYGTYLLWELQAPTGYVQLTSPVSFTIGKDKTRKELVTVVKNNKRPRIDVDPDTGE
 ETLYILMLVAILLFGSGYYLTKKTNN

ORF6_23FTW is a putative sortase. An example of an amino acid sequence of ORF6_23FTW is set forth in SEQ ID NO: 233.

SEQ ID NO: 233

MLIKMVKTKKQKRNNLLGVVFFIGMAVMAYPLVSRLYYRVESNQIADFDKEKATLDEADI DERMKLAQAFNDS
 LNNVVS GPDWSEEMKKKGRAEYARMLEIHERMGHVEIPVIDVDLPVYAGTAEVLQOGAGQLEGTS LPIGKNSTH
 AVITAHTGLPTAKMFTDLTKLVGDKFYVHNIKEVMAYQVDQVKVIEPTNFDDLLIVPGHDYVTLTCTPYMINT
 HRLLVGRHRI PYVAEVEEFIAANKLSHLRYLYFVAVGLIVILLWII RRLRKKKKQPEKALKALKAAARKEVKVE
 DGQQ

ORF7_23FTW is a putative sortase. An example of an amino acid sequence of ORF7_23FTW is set forth in SEQ ID NO: 234.

SEQ ID NO: 234

MDNSRRSRKKGTKKKKHPLILLI FLVGFVAIYPLVSRYYYRIESNEVIKEFDETVS QMDKAELEERWRLAQAF
 NATLKPSEILD PFTEQEKKKG VSEYANMLKVHERIGYVEIPAI DQEI PMYVGTSEEILQKGAGLLEGASLPVGGE
 NTHTVVTAH RGLPTAELFSQLDKMKKGDV FYLHVLDQVLAYQVDQILTVEPNDFEPVLIQH GKDYATLLTCTPYM
 INSHRLLV RGKRI PYTAPIAERNRAVRER GQFWLWLLAALVMILVLSYG VYRHRIRIVKGLEKQLEEHVKG

ORF8_23FTW is a putative sortase. An example of an amino acid sequence of ORF8_23FTW is set forth in SEQ ID NO: 235.

SEQ ID NO: 235

MSKAKLQKLLGYLLMLVALVIPVYCFGQMVLSGLQVKGHEIFSESVTADSYQEQLQSRSLDYNQRLDSQNRIVDP
 FLAEGYEVNYQVSDDDPAVYGYLSIPSLEIMEPVYLGADYHHLAMGLAHVDGTPLPVEGKGIRSVIAGHRAEPSH
 VFFRHLDQLKVG DALYYDNGQEIVEYQMMDETIILPSEWEKLESVSSKNIMTLITCDPIPTFNKRLLVNFERVAV
 YQKSDPQTA AAVARVAFTKEGQSVSRVATSQWLYRGLVVLAF LGILFVLWKLARLLRGK

As discussed above, a *S. pneumoniae* AI sequence is present in the 23F Poland 16 *S. pneumoniae* genome. Examples of *S. pneumoniae* AI sequences from 23F Poland 16 are set forth below.

ORF2_23FP is a transcriptional regulator. An example of an amino acid sequence of ORF2_23FP is set forth in SEQ ID NO: 236.

SEQ ID NO: 236

MLNKYIEKRITDKITILNILLDIRSIELDELSTLTSLQSKSLLSILQELQETFEELTFNLD TQQVQLIEHHSQ
 TNYFFHQLYNQSTILKILRFFLLQGNQSFNEFTQKEYIS IATGYRVRQKCGLLRSVGLDLVKNQVVGPEYRIRF
 LIALQLQFHFGIEIYDLNDGSM DWVTHMIVQSNSQLSHELLEITPDEYVHFSILVALTWKRREFPLEFPESKEFEK
 LKNLFMYPI LMEHCQTYLEPHANMTFTQEELDYIFLVYCSANSSFSKDKWNQEKKTHTIQILQHTRGKHL LSKF
 KNILGNDISNSLSFLTALTFLTRTFLFGLQNLVPYNYEYHYGIESDKPLYHISKAI VQEWMTQEKIEGVIDQHR
 LYLFSLYLTETIFSSLP AIPIFIILNNQADVNLIKSIILRNFTDKVASVTGYNILISPPPSEEHLTEPLIIITTK
 EYLPYVKKQY PKGKHFLTIALDLHVSQQRLIYQTI VDIRKEAFDKRVAMI AKKAHYLL

ORF3_23FP is a cell wall surface protein. An example of an amino acid sequence of ORF3_23FP is set forth in SEQ ID NO: 237.

SEQ ID NO: 237

MKKVRKLEQKAVATLCTTSQITAFSSSLVALAEETPETSIPAIGKVVIKETGEGGALLGDAVFELKNNTDGTTSVQRT
 EAQTGEAIFSNIKPGTYTLTEAQQPPVGYKPKSTKQWTVVEVEKNGRTTVQGEQVENREEALSDQYPQTGTYPDVQTP
 YQIIKVDGSEKNGQHKALNPYPYERVIPEGTLISKRIYQVNNLDDNQYGIELTVSGKTTVETKEASTPLDVVILLD
 NSNSMSNIRHNHHAHRAEKAGEATRALVDKITSNPDNRVALVTYGSTIFDGSEATVEKGVADANGKILNDSALWTF
 5 DRTTFTAKTYNYSFLNLTSDPTDIQTIKDRIPSDAEELNKDKLMTYQFGATFTQKALMTADDILTQARPNSKKVI
 FHITDGVPTMSYPINEKYTGTTQSYRTQLNNFKAKTPNSSGILLEDFVTWSADGEHKIVRGDGESYQMFTKKPVT
 DQYGVHQILSITSMEQRAKLVSAGYRFYGTDLVLYWRDSILAYPFNSSTDWITNHGDPPTWYNGNMAQDGYDVF
 TVGVGVNGDPGTDEATATRFMQSISSSPDNNTNVADPSQILQELNRYFYTIVNEKKSIENTITDPMGELIDFQL
 10 GADGRFDPADYTLTANDGSSLVNNVPTGGPQNDGGLLKNAKVFDYDTEKRIRVTGLYLGTGEKVTLTYNVRLNDQ
 FVSNKFYDTNGRTTLHPKEVEKNTVRDFPIPKIRDVRKYPEITIPKEKKLGEIEFIKINKNDKKPLRDAVFSLOK
 QHPDYPDIYGAIDQNGTYQNVRTGEDGKLTFFKNLSDGKYRLFENSEPAGYKPVQNKPIVAFQIVNGEVRDVTISIV
 PQDIPAGYEFTNDKHYITNEPIPPKREYPRTGIGMLPFYLGICMMMGVLLYTRKNP

ORF4_23FP is a cell wall surface protein. An example of an amino acid sequence of

ORF4_23FP is set forth in SEQ ID NO: 238.

SEQ ID NO: 238

MKSINKFLTMLAALLLTASSLFSAAATVFAADNVSTAPDAVTKTLTIHKLLLSDDLKTWDTNGPKGYDGTQSSLK
 DLTGVVAEEIPNVYFELQKYNLTLDGKEKENLKDDSKWTTVHGGLTTKDGLKIETSTLKGVYRIRREDRTKTTYVGP
 NGQVLTGSKAVPALVTLPLVNNNGTVIDAHVFPKNSYNKPVVDKRIADTLNNDQNGLSIGTKIPYVVNTTIPSN
 15 ATATSFWSDEMTEGLTYNEDVTITLNNVAMDQADYEVTKGINGFNLKLTEAGLAKINGKDADQKIQITYSATLN
 20 SLAVADIPESNDITYHYGNHQDHGNTPKPTKPNNGQITVTKTWDSQPAPEGVKATVQLVNAKTGEKVGAPELSE
 NNWYTWSGLDNSIEYKVEEYNGYSAEYTVESKGLGVKNWKNNPAPINLEEPRVKTYGKKFVKVDQKDTRLE
 NAQFVVKKADSNKYIAFKSTAQQAADEKAAATAKQKLDAAVAAYTNAADKQAAQALVDQAQQEYNVAYKEAKFGY
 VEVAGKDEAMVLTSNTDQGFQISGLAAGTYKLEEIKAPEGFAKIDDFEVVVGAGSWNQGEFNYLKDVQKNDATKV
 VNKKITIPQTGGIGITIFAVAGAVIMGIAVYAYVKNKDEDQLA

ORF5_23FP is a cell wall surface protein. An example of an amino acid sequence of

ORF5_23FP is set forth in SEQ ID NO: 239.

SEQ ID NO: 239

MTMQKMQKMISRIFFVMALCFSLVWGAHAVQAQEDHTLVQLQENYQEVVSQPSRDGHRQLQVWKLLDSSYSYDNRV
 30 QIVRDLHSWDENKLSSEFKKTSFEMTFLENQIEVSHIPNGLYYVRSIIQTDAVSYPAEFLFEMTDQTVPEPLVIVAK
 KADTVTTKVKLIKVDQDHNRLGEGVFKLVSVARDGSEKEVPLIGEYRYSSTGQVGRITLYTDKNGEIVVTNLPLGT
 YRFKEVEPLAGYAVTMDTDVQLVDHQLVTITVVNQKLPNGNVDFMKVDGRTNTSLQGAMFKVMKEENGHYTPVL
 QNGKEVVVASGKDGFRFRVEGLEYGTYLWELQAPTGYVQLTSPVSFTIGKDRKELVTVVKNKRPRIDVPDTGE
 35 ETLYILMLVAILLFGSGYYLTCKTNN

ORF6_23FP is a putative sortase. An example of an amino acid sequence of ORF6_23FP is
 set forth in SEQ ID NO: 240.

SEQ ID NO: 240

MLIKMAKTKKQKRNNLLGVVFFIGIAVMAYPLVSRLYRVESNQIADFDKEKATLDEADIDERMKLAQAFNDS
 40 LNNVSGDPWSEEMKKKGRAEYARMLEIHERMGHVEIPAIQVDLPVYAGTAEVLQQGAGHLEGTSLPIGNSTH
 AVITAHTGLPTAKMFTDLTKLVGDKFYVHNIKEVMAYQVDQVKVIEPTNFDDLLIVPGHDYVTLTCTPYMINT
 HRLVLRGHRIPYVAEVEEFIAANKLSHLYRYLFYVAVGLIVILLWIIIRLRKKKRQSERALKALKEATKEVKVE
 DE

ORF7_23FP is a putative sortase. An example of an amino acid sequence of ORF7_23FP is
 set forth in SEQ ID NO: 241.

SEQ ID NO: 241

MSKSRYSRKKSVKKKNPFILLIFLVGLAVAMYPLVSRYRYRIESNEVIKEFDETVSQMDKAELEERWRLAQAF
 50 NATLKPSEILDPTFEQEKKKGVSEYANMLKVHERIGYVEIPAIQDQEIIPMYVGTSEEILQKGAGLLEGASLPVGGE
 NTHTVTAHRGLPTAELFSQLDKMKKGDI FYLHVLDQVLAYQVDQIVTVEPNDFEPVLIQHGEDIATLLTCTPYM
 INSHRLVLRGKRIPYTAPIAERNRAVRERQGFWLWLLLGAMAVILLLLYRVYRNRIRIVKGLEKQLEGRHVKD

ORF8_23FP is a putative sortase. An example of an amino acid sequence of ORF8_23FP is
 set forth in SEQ ID NO: 242.

SEQ ID NO: 143 05 / 27 239

MSRTKLRLALLGYLLMLVACLIPIYCFGQMVLQSLGQVKGHATFVKSMTTTEMYQEQQNHSLAYNQRLASQNRIVDP
FLAEGYEVNYQVSDDDPAVYGYLSIPSLIMEPVYLGADYHHLGMGLAHVDGTPLPLDGTGIRSVIAGHRAEPSH
VFFRHLQDLKVGDALYYDNGQEIVEYQMMDEIILPSEWEKLESVSSKNIMTLITCDPIPTFNKRLLVNFERVAV
YQKSDPQTAAVARVAFTKEGQSVSRVATSQWLYRGLVVLAFILGILFVLWKLARLLRGK

Immunogenic compositions of the invention comprising AI antigens may further comprise one or more antigenic agents. Preferred antigens include those listed below. Additionally, the compositions of the present invention may be used to treat or prevent infections caused by any of the below-listed microbes. Antigens for use in the immunogenic compositions include, but are not limited to, one or more of the following set forth below, or antigens derived from one or more of the following set forth below:

Bacterial Antigens

N. meningitides: a protein antigen from *N. meningitides* serogroup A, C, W135, Y, and/or B (1-7); an outer-membrane vesicle (OMV) preparation from *N. meningitides* serogroup B. (8, 9, 10, 11); a saccharide antigen, including LPS, from *N. meningitides* serogroup A, B, C W135 and/or Y, such as the oligosaccharide from serogroup C (see PCT/US99/09346; PCT IB98/01665; and PCT IB99/00103);

Streptococcus pneumoniae: a saccharide or protein antigen, particularly a saccharide from *Streptococcus pneumoniae*;

Streptococcus agalactiae: particularly, Group B streptococcus antigens;

Streptococcus pyogenes: particularly, Group A streptococcus antigens;

Enterococcus faecalis or *Enterococcus faecium*: Particularly a trisaccharide repeat or other *Enterococcus* derived antigens provided in US Patent No. 6,756,361;

Helicobacter pylori: including: Cag, Vac, Nap, HopX, HopY and/or urease antigen;

Bordetella pertussis: such as pertussis holotoxin (PT) and filamentous haemagglutinin (FHA) from *B. pertussis*, optionally also combination with pertactin and/or agglutinogens 2 and 3 antigen;

Staphylococcus aureus: including *S. aureus* type 5 and 8 capsular polysaccharides optionally conjugated to nontoxic recombinant *Pseudomonas aeruginosa* exotoxin A, such as StaphVAX™, or antigens derived from surface proteins, invasins (leukocidin, kinases, hyaluronidase), surface factors that inhibit phagocytic engulfment (capsule, Protein A), carotenoids, catalase production, Protein A, coagulase, clotting factor, and/or membrane-damaging toxins (optionally detoxified) that lyse eukaryotic cell membranes (hemolysins, leukotoxin, leukocidin);

Staphylococcus epidermis: particularly, *S. epidermidis* slime-associated antigen (SAA);

Staphylococcus saprophyticus: (causing urinary tract infections) particularly the 160 kDa hemagglutinin of *S. saprophyticus* antigen;

Pseudomonas aeruginosa: particularly, endotoxin A, Wzz protein, *P. aeruginosa* LPS, more particularly LPS isolated from PAO1 (O5 serotype), and/or Outer Membrane Proteins, including Outer Membrane Proteins F (OprF) (*Infect Immun.* 2001 May; 69(5): 3510-3515);

~~Bacillus anthracis~~ (anthrax): such as *B. anthracis* antigens (optionally detoxified) from A-components (lethal factor (LF) and edema factor (EF)), both of which can share a common B-component known as protective antigen (PA);

Moraxella catarrhalis: (respiratory) including outer membrane protein antigens (HMW-OMP), C-antigen, and/or LPS;

Yersinia pestis (plague): such as F1 capsular antigen (*Infect Immun.* 2003 Jan; 71(1)): 374-383, LPS (*Infect Immun.* 1999 Oct; 67(10): 5395), *Yersinia pestis* V antigen (*Infect Immun.* 1997 Nov; 65(11): 4476-4482);

Yersinia enterocolitica (gastrointestinal pathogen): particularly LPS (*Infect Immun.* 2002 August; 70(8): 4414);

Yersinia pseudotuberculosis: gastrointestinal pathogen antigens;

Mycobacterium tuberculosis: such as lipoproteins, LPS, BCG antigens, a fusion protein of antigen 85B (Ag85B) and/or ESAT-6 optionally formulated in cationic lipid vesicles (*Infect Immun.* 2004 October; 72(10): 6148), *Mycobacterium tuberculosis* (Mtb) isocitrate dehydrogenase associated antigens (*Proc Natl Acad Sci U S A.* 2004 Aug 24; 101(34): 12652), and/or MPT51 antigens (*Infect Immun.* 2004 July; 72(7): 3829);

Legionella pneumophila (Legionnaires' Disease): *L. pneumophila* antigens -- optionally derived from cell lines with disrupted *asd* genes (*Infect Immun.* 1998 May; 66(5): 1898);

Rickettsia: including outer membrane proteins, including the outer membrane protein A and/or B (OmpB) (*Biochim Biophys Acta.* 2004 Nov 1; 1702(2):145), LPS, and surface protein antigen (SPA) (*J Autoimmun.* 1989 Jun; 2 Suppl:81);

E. coli: including antigens from enterotoxigenic *E. coli* (ETEC), enteroaggregative *E. coli* (EAaggEC), diffusely adhering *E. coli* (DAEC), enteropathogenic *E. coli* (EPEC), and/or enterohemorrhagic *E. coli* (EHEC);

Vibrio cholerae: including proteinase antigens, LPS, particularly lipopolysaccharides of *Vibrio cholerae* II, O1 Inaba O-specific polysaccharides, *V. cholera* O139, antigens of IEM108 vaccine (*Infect Immun.* 2003 Oct; 71(10):5498-504), and/or Zonula occludens toxin (Zot);

Salmonella typhi (typhoid fever): including capsular polysaccharides preferably conjugates (Vi, i.e. vax-TyVi);

Salmonella typhimurium (gastroenteritis): antigens derived therefrom are contemplated for microbial and cancer therapies, including angiogenesis inhibition and modulation of flk;

Listeria monocytogenes (systemic infections in immunocompromised or elderly people, infections of fetus): antigens derived from *L. monocytogenes* are preferably used as carriers/vectors for intracytoplasmic delivery of conjugates/associated compositions of the present invention;

Porphyromonas gingivalis: particularly, *P. gingivalis* outer membrane protein (OMP);

Tetanus: such as tetanus toxoid (TT) antigens, preferably used as a carrier protein in conjunction/conjugated with the compositions of the present invention;

~~For Diphtheria~~ such as a diphtheria toxoid, preferably CRM₁₉₇, additionally antigens capable of modulating, inhibiting or associated with ADP ribosylation are contemplated for combination/co-administration/conjugation with the compositions of the present invention, the diphtheria toxoids are preferably used as carrier proteins;

5 *Borrelia burgdorferi* (Lyme disease): such as antigens associated with P39 and P13 (an integral membrane protein, *Infect Immun.* 2001 May; 69(5): 3323-3334), VlsE Antigenic Variation Protein (*J Clin Microbiol.* 1999 Dec; 37(12): 3997);

Haemophilus influenzae B: such as a saccharide antigen therefrom;

10 *Klebsiella*: such as an OMP, including OMP A, or a polysaccharide optionally conjugated to tetanus toxoid;

Neisseria gonorrhoeae: including, a Por (or porin) protein, such as PorB (*see Zhu et al., Vaccine* (2004) 22:660 – 669), a transferring binding protein, such as TbpA and TbpB (*See Price et al., Infection and Immunity* (2004) 71(1):277 – 283), a opacity protein (such as Opa), a reduction-modifiable protein (Rmp), and outer membrane vesicle (OMV) preparations (*see Plante et al., J Infectious Disease* (2000) 182:848 – 855), also see *e.g.* WO99/24578, WO99/36544, WO99/57280, WO02/079243);

Chlamydia pneumoniae: particularly *C. pneumoniae* protein antigens;

20 *Chlamydia trachomatis*: including antigens derived from serotypes A, B, Ba and C are (agents of trachoma, a cause of blindness), serotypes L₁, L₂ & L₃ (associated with Lymphogranuloma venereum), and serotypes, D-K;

Treponema pallidum (Syphilis): particularly a TmpA antigen; and

Haemophilus ducreyi (causing chancroid): including outer membrane protein (DsrA).

Where not specifically referenced, further bacterial antigens of the invention may be capsular antigens, polysaccharide antigens or protein antigens of any of the above. Further bacterial antigens
25 may also include an outer membrane vesicle (OMV) preparation. Additionally, antigens include live, attenuated, split, and/or purified versions of any of the aforementioned bacteria. The bacterial or microbial derived antigens of the present invention may be gram-negative or gram-positive and aerobic or anaerobic.

30 Additionally, any of the above bacterial-derived saccharides (polysaccharides, LPS, LOS or oligosaccharides) can be conjugated to another agent or antigen, such as a carrier protein (for example CRM₁₉₇). Such conjugation may be direct conjugation effected by reductive amination of carbonyl moieties on the saccharide to amino groups on the protein, as provided in US Patent No. 5,360,897 and *Can J Biochem Cell Biol.* 1984 May;62(5):270-5. Alternatively, the saccharides can be conjugated through a linker, such as, with succinamide or other linkages provided in *Bioconjugate Techniques*, 1996 and *CRC, Chemistry of Protein Conjugation and Cross-Linking*, 1993.

Poliovirus

Influenza: including whole viral particles (attenuated), split, or subunit comprising hemagglutinin (HA) and/or neuraminidase (NA) surface proteins, the influenza antigens may be derived from chicken embryos or propagated on cell culture, and/or the influenza antigens are derived from influenza type A, B, and/or C, among others;

Respiratory syncytial virus (RSV): including the F protein of the A2 strain of RSV (*J Gen Virol.* 2004 Nov; 85(Pt 11):3229) and/or G glycoprotein;

Parainfluenza virus (PIV): including PIV type 1, 2, and 3, preferably containing hemagglutinin, neuraminidase and/or fusion glycoproteins;

Poliovirus: including antigens from a family of picornaviridae, preferably poliovirus antigens such as OPV or, preferably IPV;

Measles: including split measles virus (MV) antigen optionally combined with the Protollin and or antigens present in MMR vaccine;

Mumps: including antigens present in MMR vaccine;

Rubella: including antigens present in MMR vaccine as well as other antigens from Togaviridae, including dengue virus;

Rabies: such as lyophilized inactivated virus (RabAvert™);

Flaviviridae viruses: such as (and antigens derived therefrom) yellow fever virus, Japanese encephalitis virus, dengue virus (types 1, 2, 3, or 4), tick borne encephalitis virus, and West Nile virus;

Caliciviridae; antigens therefrom;

HIV: including HIV-1 or HIV-2 strain antigens, such as gag (p24gag and p55gag), env (gp160 and gp41), pol, tat, nef, rev vpu, miniproteins, (preferably p55 gag and gp140v delete) and antigens from the isolates HIV_{IIIb}, HIV_{SF2}, HIV_{LAV}, HIV_{LAI}, HIV_{MN}, HIV-1_{CM235}, HIV-1_{US4}, HIV-2; simian immunodeficiency virus (SIV) among others;

Rotavirus: including VP4, VP5, VP6, VP7, VP8 proteins (*Protein Expr Purif.* 2004 Dec;38(2):205) and/or NSP4;

Pestivirus: such as antigens from classical porcine fever virus, bovine viral diarrhoea virus, and/or border disease virus;

Parvovirus: such as parvovirus B19;

Coronavirus: including SARS virus antigens, particularly spike protein or proteases therefrom, as well as antigens included in WO 04/92360;

Hepatitis A virus: such as inactivated virus;

Hepatitis B virus: such as the surface and/or core antigens (sAg), as well as the presurface sequences, pre-S1 and pre-S2 (formerly called pre-S), as well as combinations of the above, such as sAg/pre-S1, sAg/pre-S2, sAg/pre-S1/pre-S2, and pre-S1/pre-S2, (see, e.g., AHBV Vaccines - *Human Vaccines and Vaccination*, pp. 159-176; and U.S. Patent Nos. 4,722,840, 5,098,704, 5,324,513;

Beaunes et al., *J. Virol.* (1995) 69:6833-6838, Birnbaum et al., *J. Virol.* (1990) 64:3319-3330; and Zhou et al., *J. Virol.* (1991) 65:5457-5464);

Hepatitis C virus: such as E1, E2, E1/E2 (see, Houghton et al., *Hepatology* (1991) 14:381), NS345 polypeptide, NS 345-core polypeptide, core, and/or peptides from the nonstructural regions (International Publication Nos. WO 89/04669; WO 90/11089; and WO 90/14436);

Delta hepatitis virus (HDV): antigens derived therefrom, particularly δ -antigen from HDV (see, e.g., U.S. Patent No. 5,378,814);

Hepatitis E virus (HEV); antigens derived therefrom;

Hepatitis G virus (HGV); antigens derived therefrom;

Varicella zoster virus: antigens derived from varicella zoster virus (VZV) (*J. Gen. Virol.* (1986) 67:1759);

Epstein-Barr virus: antigens derived from EBV (Baer et al., *Nature* (1984) 310:207);

Cytomegalovirus: CMV antigens, including gB and gH (*Cytomegaloviruses* (J.K. McDougall, ed., Springer-Verlag 1990) pp. 125-169);

Herpes simplex virus: including antigens from HSV-1 or HSV-2 strains and glycoproteins gB, gD and gH (McGeoch et al., *J. Gen. Virol.* (1988) 69:1531 and U.S. Patent No. 5,171,568);

Human Herpes Virus: antigens derived from other human herpesviruses such as HHV6 and HHV7; and

HPV: including antigens associated with or derived from human papillomavirus (HPV), for example, one or more of E1 – E7, L1, L2, and fusions thereof, particularly the compositions of the invention may include a virus-like particle (VLP) comprising the L1 major capsid protein, more particular still, the HPV antigens are protective against one or more of HPV serotypes 6, 11, 16 and/or 18.

Further provided are antigens, compositions, methods, and microbes included in *Vaccines*, 4th Edition (Plotkin and Orenstein ed. 2004); *Medical Microbiology* 4th Edition (Murray et al. ed. 2002); *Virology*, 3rd Edition (W.K. Joklik ed. 1988); *Fundamental Virology*, 2nd Edition (B.N. Fields and D.M. Knipe, eds. 1991), which are contemplated in conjunction with the compositions of the present invention.

Additionally, antigens include live, attenuated, split, and/or purified versions of any of the aforementioned viruses.

Fungal Antigens

Fungal antigens for use herein, associated with vaccines include those described in: U.S. Pat. Nos. 4,229,434 and 4,368,191 for prophylaxis and treatment of trichophytosis caused by *Trichophyton mentagrophytes*; U.S. Pat. Nos. 5,277,904 and 5,284,652 for a broad spectrum dermatophyte vaccine for the prophylaxis of dermatophyte infection in animals, such as guinea pigs, cats, rabbits, horses and lambs, these antigens comprises a suspension of killed *T. equinum*, *T. mentagrophytes* (var. *granulare*), *M. canis* and/or *M. gypseum* in an effective amount optionally combined with an adjuvant;

U.S. Pat. Nos. 5,453,275 and 6,132,735 for a ringworm vaccine comprising an effective amount of a homogenized, formaldehyde-killed fungi, i.e., *Microsporum canis* culture in a carrier; U.S. Pat. No. 5,948,413 involving extracellular and intracellular proteins for pythiosis. Additional antigens identified within antifungal vaccines include Ringvac bovis LTF-130 and Bioveta.

Further, fungal antigens for use herein may be derived from Dermatophytes, including: *Epidermophyton floccosum*, *Microsporum audouini*, *Microsporum canis*, *Microsporum distortum*, *Microsporum equinum*, *Microsporum gypsum*, *Microsporum nanum*, *Trichophyton concentricum*, *Trichophyton equinum*, *Trichophyton gallinae*, *Trichophyton gypseum*, *Trichophyton megnini*, *Trichophyton mentagrophytes*, *Trichophyton quinckeanum*, *Trichophyton rubrum*, *Trichophyton schoenleini*, *Trichophyton tonsurans*, *Trichophyton verrucosum*, *T. verrucosum* var. album, var. discoides, var. ochraceum, *Trichophyton violaceum*, and/or *Trichophyton faviforme*.

Fungal pathogens for use as antigens or in derivation of antigens in conjunction with the compositions of the present invention comprise *Aspergillus fumigatus*, *Aspergillus flavus*, *Aspergillus niger*, *Aspergillus nidulans*, *Aspergillus terreus*, *Aspergillus sydowi*, *Aspergillus flavatus*, *Aspergillus glaucus*, *Blastoschizomyces capitatus*, *Candida albicans*, *Candida enolase*, *Candida tropicalis*, *Candida glabrata*, *Candida krusei*, *Candida parapsilosis*, *Candida stellatoidea*, *Candida kusei*, *Candida parakwsei*, *Candida lusitaniae*, *Candida pseudotropicalis*, *Candida guilliermondi*, *Cladosporium carrionii*, *Coccidioides immitis*, *Blastomyces dermatidis*, *Cryptococcus neoformans*, *Geotrichum clavatum*, *Histoplasma capsulatum*, *Klebsiella pneumoniae*, *Paracoccidioides brasiliensis*, *Pneumocystis carinii*, *Pythium insidiosum*, *Pityrosporum ovale*, *Saccharomyces cerevisiae*, *Saccharomyces boulardii*, *Saccharomyces pombe*, *Scedosporium apiospermum*, *Sporothrix schenckii*, *Trichosporon beigeli*, *Toxoplasma gondii*, *Penicillium marneffe*, *Malassezia* spp., *Fonsecaea* spp., *Wangiella* spp., *Sporothrix* spp., *Basidiobolus* spp., *Conidiobolus* spp., *Rhizopus* spp., *Mucor* spp, *Absidia* spp, *Mortierella* spp, *Cunninghamella* spp, and *Saksenaea* spp.

Other fungi from which antigens are derived include *Alternaria* spp, *Curvularia* spp, *Helminthosporium* spp, *Fusarium* spp, *Aspergillus* spp, *Penicillium* spp, *Monolinia* spp, *Rhizoctonia* spp, *Paecilomyces* spp, *Pithomyces* spp, and *Cladosporium* spp.

Processes for producing a fungal antigens are well known in the art (see US Patent No. 6,333,164). In a preferred method a solubilized fraction extracted and separated from an insoluble fraction obtainable from fungal cells of which cell wall has been substantially removed or at least partially removed, characterized in that the process comprises the steps of: obtaining living fungal cells; obtaining fungal cells of which cell wall has been substantially removed or at least partially removed; bursting the fungal cells of which cell wall has been substantially removed or at least partially removed; obtaining an insoluble fraction; and extracting and separating a solubilized fraction from the insoluble fraction.

STD Antigens

In particular embodiments, microbes (bacteria, viruses and/or fungi) against which the present compositions and methods can be implemented include those that cause sexually transmitted diseases (STDs) and/or those that display on their surface an antigen that can be the target or antigen composition of the invention. In a preferred embodiment of the invention, compositions are combined with antigens derived from a viral or bacterial STD. Antigens derived from bacteria or viruses can be administered in conjunction with the compositions of the present invention to provide protection against at least one of the following STDs, among others: chlamydia, genital herpes, hepatitis (particularly HCV), genital warts, gonorrhoea, syphilis and/or chancroid (See, WO00/15255).

In another embodiment the compositions of the present invention are co-administered with an antigen for the prevention or treatment of an STD.

Antigens derived from the following viruses associated with STDs, which are described in greater detail above, are preferred for co-administration with the compositions of the present invention: hepatitis (particularly HCV), HPV, HIV, or HSV.

Additionally, antigens derived from the following bacteria associated with STDs, which are described in greater detail above, are preferred for co-administration with the compositions of the present invention: *Neisseria gonorrhoeae*, *Chlamydia pneumoniae*, *Chlamydia trachomatis*, *Treponema pallidum*, or *Haemophilus ducreyi*.

Respiratory Antigens

The antigen may be a respiratory antigen and could further be used in an immunogenic composition for methods of preventing and/or treating infection by a respiratory pathogen, including a virus, bacteria, or fungi such as respiratory syncytial virus (RSV), PIV, SARS virus, influenza, *Bacillus anthracis*, particularly by reducing or preventing infection and/or one or more symptoms of respiratory virus infection. A composition comprising an antigen described herein, such as one derived from a respiratory virus, bacteria or fungus is administered in conjunction with the compositions of the present invention to an individual which is at risk of being exposed to that particular respiratory microbe, has been exposed to a respiratory microbe or is infected with a respiratory virus, bacteria or fungus. The composition(s) of the present invention is/are preferably co-administered at the same time or in the same formulation with an antigen of the respiratory pathogen. Administration of the composition results in reduced incidence and/or severity of one or more symptoms of respiratory infection.

Pediatric/Geriatric Antigens

In one embodiment the compositions of the present invention are used in conjunction with an antigen for treatment of a pediatric population, as in a pediatric antigen. In a more particular embodiment the pediatric population is less than about 3 years old, or less than about 2 years, or less than about 1 years old. In another embodiment the pediatric antigen (in conjunction with the composition of the present invention) is administered multiple times over at least 1, 2, or 3 years.

In another embodiment the compositions of the present invention are used in conjunction with an antigen for treatment of a geriatric population, as in a geriatric antigen.

Other Antigens

Other antigens for use in conjunction with the compositions of the present include hospital acquired (nosocomial) associated antigens.

In another embodiment, parasitic antigens are contemplated in conjunction with the compositions of the present invention. Examples of parasitic antigens include those derived from organisms causing malaria and/or Lyme disease.

In another embodiment, the antigens in conjunction with the compositions of the present invention are associated with or effective against a mosquito born illness. In another embodiment, the antigens in conjunction with the compositions of the present invention are associated with or effective against encephalitis. In another embodiment the antigens in conjunction with the compositions of the present invention are associated with or effective against an infection of the nervous system.

In another embodiment, the antigens in conjunction with the compositions of the present invention are antigens transmissible through blood or body fluids.

Antigen Formulations

In other aspects of the invention, methods of producing microparticles having adsorbed antigens are provided. The methods comprise: (a) providing an emulsion by dispersing a mixture comprising (i) water, (ii) a detergent, (iii) an organic solvent, and (iv) a biodegradable polymer selected from the group consisting of a poly(α -hydroxy acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester, a polyanhydride, and a polycyanoacrylate. The polymer is typically present in the mixture at a concentration of about 1% to about 30% relative to the organic solvent, while the detergent is typically present in the mixture at a weight-to-weight detergent-to-polymer ratio of from about 0.00001:1 to about 0.1:1 (more typically about 0.0001:1 to about 0.1:1, about 0.001:1 to about 0.1:1, or about 0.005:1 to about 0.1:1); (b) removing the organic solvent from the emulsion; and (c) adsorbing an antigen on the surface of the microparticles. In certain embodiments, the biodegradable polymer is present at a concentration of about 3% to about 10% relative to the organic solvent.

Microparticles for use herein will be formed from materials that are sterilizable, non-toxic and biodegradable. Such materials include, without limitation, poly(α -hydroxy acid), polyhydroxybutyric acid, polycaprolactone, polyorthoester, polyanhydride, PACA, and polycyanoacrylate. Preferably, microparticles for use with the present invention are derived from a poly(α -hydroxy acid), in particular, from a poly(lactide) ("PLA") or a copolymer of D,L-lactide and glycolide or glycolic acid, such as a poly(D,L-lactide-co-glycolide) ("PLG" or "PLGA"), or a copolymer of D,L-lactide and caprolactone. The microparticles may be derived from any of various polymeric starting materials which have a variety of molecular weights and, in the case of the copolymers such as PLG, a variety of lactide:glycolide ratios, the selection of which will be largely a

matter of choice, depending in part on the coadministered macromolecule. These parameters are discussed more fully below.

Further antigens may also include an outer membrane vesicle (OMV) preparation.

Additional formulation methods and antigens (especially tumor antigens) are provided in U.S.

5 Patent Serial No. 09/581,772.

Antigen References

The following references include antigens useful in conjunction with the compositions of the present invention:

- 10 1 International patent application WO99/24578
- 2 International patent application WO99/36544.
- 3 International patent application WO99/57280.
- 4 International patent application WO00/22430.
- 5 Tettelin et al. (2000) Science 287:1809-1815.
- 15 6 International patent application WO96/29412.
- 7 Pizza et al. (2000) Science 287:1816-1820.
- 8 PCT WO 01/52885.
- 9 Bjune et al. (1991) Lancet 338(8775).
- 10 Fuskasawa et al. (1999) Vaccine 17:2951-2958.
- 20 11 Rosenqvist et al. (1998) Dev. Biol. Stand 92:323-333.
- 12 Constantino et al. (1992) Vaccine 10:691-698.
- 13 Constantino et al. (1999) Vaccine 17:1251-1263.
- 14 Watson (2000) Pediatr Infect Dis J 19:331-332.
- 15 Rubin (2000) Pediatr Clin North Am 47:269-285, v.
- 25 16 Jedrzejewski (2001) Microbiol Mol Biol Rev 65:187-207.
- 17 International patent application filed on 3rd July 2001 claiming priority from GB-0016363.4; WO 02/02606; PCT IB/01/00166.
- 18 Kalman et al. (1999) Nature Genetics 21:385-389.
- 19 Read et al. (2000) Nucleic Acids Res 28:1397-406.
- 30 20 Shirai et al. (2000) J. Infect. Dis 181(Suppl 3):S524-S527.
- 21 International patent application WO99/27105.
- 22 International patent application WO00/27994.
- 23 International patent application WO00/37494.
- 24 International patent application WO99/28475.
- 35 25 Bell (2000) Pediatr Infect Dis J 19:1187-1188.
- 26 Iwarson (1995) APMIS 103:321-326.
- 27 Gerlich et al. (1990) Vaccine 8 Suppl:S63-68 & 79-80.
- 28 Hsu et al. (1999) Clin Liver Dis 3:901-915.
- 29 Gastofsson et al. (1996) N. Engl. J. Med. 334:349-355.
- 40 30 Rappuoli et al. (1991) TIBTECH 9:232-238.
- 31 Vaccines (1988) eds. Plotkin & Mortimer. ISBN 0-7216-1946-0.
- 32 Del Giudice et al. (1998) Molecular Aspects of Medicine 19:1-70.
- 33 International patent application WO93/018150.
- 34 International patent application WO99/53310.
- 45 35 International patent application WO98/04702.
- 36 Ross et al. (2001) Vaccine 19:135-142.
- 37 Sutter et al. (2000) Pediatr Clin North Am 47:287-308.
- 38 Zimmerman & Spann (1999) Am Fam Physician 59:113-118, 125-126.
- 39 Dreensen (1997) Vaccine 15 Suppl:S2-6.
- 50 40 MMWR Morb Mortal Wkly rep 1998 Jan 16:47(1):12, 9.
- 41 McMichael (2000) Vaccine 19 Suppl 1:S101-107.

- 42 Schuchat (1999) *Lancet* 353(9146):51-6.
- 43 GB patent applications 0026333.5, 0028727.6 & 0105640.7.
- 44 Dale (1999) *Infect Disclin North Am* 13:227-43, viii.
- 45 Ferretti et al. (2001) *PNAS USA* 98: 4658-4663.
- 5 46 Kuroda et al. (2001) *Lancet* 357(9264):1225-1240; see also pages 1218-1219.
- 47 Ramsay et al. (2001) *Lancet* 357(9251):195-196.
- 48 Lindberg (1999) *Vaccine* 17 Suppl.2:S28-36.
- 49 Buttery & Moxon (2000) *J R Coil Physicians Long* 34:163-168.
- 50 Ahmad & Chapnick (1999) *Infect Dis Clin North Am* 13:113-133, vii.
- 10 51 Goldblatt (1998) *J. Med. Microbiol.* 47:663-567.
- 52 European patent 0 477 508.
- 53 U.S. Patent No. 5,306,492.
- 54 International patent application WO98/42721.
- 55 Conjugate Vaccines (eds. Cruse et al.) ISBN 3805549326, particularly vol. 10:48-114.
- 15 56 Hermanson (1996) *Bioconjugate Techniques* ISBN: 012323368 & 012342335X.
- 57 European patent application 0372501.
- 58 European patent application 0378881.
- 59 European patent application 0427347.
- 60 International patent application WO93/17712.
- 20 61 International patent application WO98/58668.
- 62 European patent application 0471177.
- 63 International patent application WO00/56360.
- 64 International patent application WO00/67161.

25 The contents of all of the above cited patents, patent applications and journal articles are incorporated by reference as if set forth fully herein.

There may be an upper limit to the number of Gram positive bacterial proteins which will be in the compositions of the invention. Preferably, the number of Gram positive bacterial proteins in a composition of the invention is less than 20, less than 19, less than 18, less than 17, less than 16, less than 15, less than 14, less than 13, less than 12, less than 11, less than 10, less than 9, less than 8, less than 7, less than 6, less than 5, less than 4, or less than 3. Still more preferably, the number of Gram positive bacterial proteins in a composition of the invention is less than 6, less than 5, or less than 4. Still more preferably, the number of Gram positive bacterial proteins in a composition of the invention is 3.

35 The Gram positive bacterial proteins and polynucleotides used in the invention are preferably isolated, *i.e.*, separate and discrete, from the whole organism with which the molecule is found in nature or, when the polynucleotide or polypeptide is not found in nature, is sufficiently free of other biological macromolecules so that the polynucleotide or polypeptide can be used for its intended purpose.

40 Fusion Proteins: GBS AI sequences

The GBS AI proteins used in the invention may be present in the composition as individual separate polypeptides, but it is preferred that at least two (*i.e.* 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, or 18) of the antigens are expressed as a single polypeptide chain (a "hybrid" or "fusion" polypeptide). Such fusion polypeptides offer two principal advantages: first, a polypeptide that may be unstable or poorly expressed on its own can be assisted by adding a suitable fusion partner that

overcomes the problem: second, commercial manufacture is simplified as only one expression and purification need be employed in order to produce two polypeptides which are both antigenically useful.

The fusion polypeptide may comprise one or more AI polypeptide sequences. Preferably, the fusion comprises an AI surface protein sequence. Preferably, the fusion polypeptide includes one or more of GBS 80, GBS 104, and GBS 67. Most preferably, the fusion peptide includes a polypeptide sequence from GBS 80. Accordingly, the invention includes a fusion peptide comprising a first amino acid sequence and a second amino acid sequence, wherein said first and second amino acid sequences are selected from a GBS AI surface protein or a fragment thereof. Preferably, the first and second amino acid sequences in the fusion polypeptide comprise different epitopes.

Hybrids (or fusions) consisting of amino acid sequences from two, three, four, five, six, seven, eight, nine, or ten GBS antigens are preferred. In particular, hybrids consisting of amino acid sequences from two, three, four, or five GBS antigens are preferred.

Different hybrid polypeptides may be mixed together in a single formulation. Within such combinations, a GBS antigen may be present in more than one hybrid polypeptide and/or as a non-hybrid polypeptide. It is preferred, however, that an antigen is present either as a hybrid or as a non-hybrid, but not as both.

Hybrid polypeptides can be represented by the formula $\text{NH}_2\text{-A-}\{-\text{X-L-}\}_n\text{-B-COOH}$, wherein: X is an amino acid sequence of a GBS AI protein or a fragment thereof; L is an optional linker amino acid sequence; A is an optional N-terminal amino acid sequence; B is an optional C-terminal amino acid sequence; and n is 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15.

If a -X- moiety has a leader peptide sequence in its wild-type form, this may be included or omitted in the hybrid protein. In some embodiments, the leader peptides will be deleted except for that of the -X- moiety located at the N-terminus of the hybrid protein *i.e.* the leader peptide of X_1 will be retained, but the leader peptides of $X_2 \dots X_n$ will be omitted. This is equivalent to deleting all leader peptides and using the leader peptide of X_1 as moiety -A-.

For each n instances of $\{-\text{X-L-}\}$, linker amino acid sequence -L- may be present or absent. For instance, when $n=2$ the hybrid may be $\text{NH}_2\text{-X}_1\text{-L}_1\text{-X}_2\text{-L}_2\text{-COOH}$, $\text{NH}_2\text{-X}_1\text{-X}_2\text{-COOH}$, $\text{NH}_2\text{-X}_1\text{-L}_1\text{-X}_2\text{-COOH}$, $\text{NH}_2\text{-X}_1\text{-X}_2\text{-L}_2\text{-COOH}$, *etc.* Linker amino acid sequence(s) -L- will typically be short (*e.g.* 20 or fewer amino acids *i.e.* 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1). Examples comprise short peptide sequences which facilitate cloning, poly-glycine linkers (*i.e.* comprising Gly_n where $n = 2, 3, 4, 5, 6, 7, 8, 9, 10$ or more), and histidine tags (*i.e.* His_n where $n = 3, 4, 5, 6, 7, 8, 9, 10$ or more). Other suitable linker amino acid sequences will be apparent to those skilled in the art. A useful linker is GSGGGG, with the Gly-Ser dipeptide being formed from a *Bam*HI restriction site, thus aiding cloning and manipulation, and the $(\text{Gly})_4$ tetrapeptide being a typical poly-glycine linker.

-A- is an optional N-terminal amino acid sequence. This will typically be short (*e.g.* 40 or fewer amino acids *i.e.* 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19,

18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1). Examples include leader sequences to direct protein trafficking, or short peptide sequences which facilitate cloning or purification (e.g. histidine tags *i.e.* His_n where *n* = 3, 4, 5, 6, 7, 8, 9, 10 or more). Other suitable N-terminal amino acid sequences will be apparent to those skilled in the art. If X₁ lacks its own N-terminus methionine, -A- is preferably an oligopeptide (e.g. with 1, 2, 3, 4, 5, 6, 7 or 8 amino acids) which provides a N-terminus methionine.

-B- is an optional C-terminal amino acid sequence. This will typically be short (e.g. 40 or fewer amino acids *i.e.* 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1). Examples include sequences to direct protein trafficking, short peptide sequences which facilitate cloning or purification (e.g. comprising histidine tags *i.e.* His_n where *n* = 3, 4, 5, 6, 7, 8, 9, 10 or more), or sequences which enhance protein stability. Other suitable C-terminal amino acid sequences will be apparent to those skilled in the art.

Most preferably, *n* is 2 or 3.

Fusion Proteins: Gram positive bacteria AI sequences

The Gram positive bacteria AI proteins used in the invention may be present in the composition as individual separate polypeptides, but it is preferred that at least two (*i.e.* 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, or 18) of the antigens are expressed as a single polypeptide chain (a "hybrid" or "fusion" polypeptide). Such fusion polypeptides offer two principal advantages: first, a polypeptide that may be unstable or poorly expressed on its own can be assisted by adding a suitable fusion partner that overcomes the problem; second, commercial manufacture is simplified as only one expression and purification need be employed in order to produce two polypeptides which are both antigenically useful.

The fusion polypeptide may comprise one or more AI polypeptide sequences. Preferably, the fusion comprises an AI surface protein sequence. Accordingly, the invention includes a fusion peptide comprising a first amino acid sequence and a second amino acid sequence, wherein said first and second amino acid sequences are selected from a Gram positive bacteria AI protein or a fragment thereof. Preferably, the first and second amino acid sequences in the fusion polypeptide comprise different epitopes.

Hybrids (or fusions) consisting of amino acid sequences from two, three, four, five, six, seven, eight, nine, or ten Gram positive bacteria antigens are preferred. In particular, hybrids consisting of amino acid sequences from two, three, four, or five Gram positive bacteria antigens are preferred.

Different hybrid polypeptides may be mixed together in a single formulation. Within such combinations, a Gram positive bacteria AI sequence may be present in more than one hybrid polypeptide and/or as a non-hybrid polypeptide. It is preferred, however, that an antigen is present either as a hybrid or as a non-hybrid, but not as both.

Hybrid polypeptides can be represented by the formula NH₂-A-{-X-L-}_n-B-COOH, wherein: X is an amino acid sequence of a Gram positive bacteria AI sequence or a fragment thereof; L is an

optional linker amino acid sequence; A is an optional N-terminal amino acid sequence; B is an optional C-terminal amino acid sequence; and n is 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15.

If a -X- moiety has a leader peptide sequence in its wild-type form, this may be included or omitted in the hybrid protein. In some embodiments, the leader peptides will be deleted except for that of the -X- moiety located at the N-terminus of the hybrid protein *i.e.* the leader peptide of X_1 will be retained, but the leader peptides of $X_2 \dots X_n$ will be omitted. This is equivalent to deleting all leader peptides and using the leader peptide of X_1 as moiety -A-.

For each n instances of {-X-L-}, linker amino acid sequence -L- may be present or absent. For instance, when $n=2$ the hybrid may be $\text{NH}_2\text{-X}_1\text{-L}_1\text{-X}_2\text{-L}_2\text{-COOH}$, $\text{NH}_2\text{-X}_1\text{-X}_2\text{-COOH}$, $\text{NH}_2\text{-X}_1\text{-L}_1\text{-X}_2\text{-COOH}$, $\text{NH}_2\text{-X}_1\text{-X}_2\text{-L}_2\text{-COOH}$, *etc.* Linker amino acid sequence(s) -L- will typically be short (*e.g.* 20 or fewer amino acids *i.e.* 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1). Examples comprise short peptide sequences which facilitate cloning, poly-glycine linkers (*i.e.* comprising Gly_n where $n = 2, 3, 4, 5, 6, 7, 8, 9, 10$ or more), and histidine tags (*i.e.* His_n where $n = 3, 4, 5, 6, 7, 8, 9, 10$ or more). Other suitable linker amino acid sequences will be apparent to those skilled in the art. A useful linker is GSGGGG, with the Gly-Ser dipeptide being formed from a *Bam*HI restriction site, thus aiding cloning and manipulation, and the $(\text{Gly})_4$ tetrapeptide being a typical poly-glycine linker.

-A- is an optional N-terminal amino acid sequence. This will typically be short (*e.g.* 40 or fewer amino acids *i.e.* 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1). Examples include leader sequences to direct protein trafficking, or short peptide sequences which facilitate cloning or purification (*e.g.* histidine tags *i.e.* His_n where $n = 3, 4, 5, 6, 7, 8, 9, 10$ or more). Other suitable N-terminal amino acid sequences will be apparent to those skilled in the art. If X_1 lacks its own N-terminus methionine, -A- is preferably an oligopeptide (*e.g.* with 1, 2, 3, 4, 5, 6, 7 or 8 amino acids) which provides a N-terminus methionine.

-B- is an optional C-terminal amino acid sequence. This will typically be short (*e.g.* 40 or fewer amino acids *i.e.* 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1). Examples include sequences to direct protein trafficking, short peptide sequences which facilitate cloning or purification (*e.g.* comprising histidine tags *i.e.* His_n where $n = 3, 4, 5, 6, 7, 8, 9, 10$ or more), or sequences which enhance protein stability. Other suitable C-terminal amino acid sequences will be apparent to those skilled in the art.

Most preferably, n is 2 or 3.

Antibodies: GBS AI sequences

The GBS AI proteins of the invention may also be used to prepare antibodies specific to the GBS AI proteins. The antibodies are preferably specific to the an oligomeric or hyper-oligomeric form of an AI protein. The invention also includes combinations of antibodies specific to GBS AI proteins selected to provide protection against an increased range of GBS serotypes and strain isolates. For example, a combination may comprise a first and second antibody, wherein said first

antibody is specific to a first GBS AI protein and said second antibody is specific to a second GBS AI protein. Preferably, the nucleic acid sequence encoding said first GBS AI protein is not present in a GBS genome comprising a polynucleotide sequence encoding for said second GBS AI protein. Preferably, the nucleic acid sequence encoding said first and second GBS AI proteins are present in the genomes of multiple GBS serotypes and strain isolates.

The GBS specific antibodies of the invention include one or more biological moieties that, through chemical or physical means, can bind to or associate with an epitope of a GBS polypeptide. The antibodies of the invention include antibodies which specifically bind to a GBS AI protein. The invention includes antibodies obtained from both polyclonal and monoclonal preparations, as well as the following: hybrid (chimeric) antibody molecules (see, for example, Winter *et al.* (1991) *Nature* 349: 293-299; and US Patent No. 4,816,567; F(ab')₂ and F(ab) fragments; F_v molecules (non-covalent heterodimers, see, for example, Inbar *et al.* (1972) *Proc Natl Acad Sci USA* 69:2659-2662; and Ehrlich *et al.* (1980) *Biochem* 19:4091-4096); single-chain Fv molecules (sFv) (see, for example, Huston *et al.* (1988) *Proc Natl Acad Sci USA* 85:5897-5883); dimeric and trimeric antibody fragment constructs; minibodies (see, *e.g.*, Pack *et al.* (1992) *Biochem* 31:1579-1584; Cumber *et al.* (1992) *J Immunology* 149B: 120-126); humanized antibody molecules (see, for example, Riechmann *et al.* (1988) *Nature* 332:323-327; Verhoeyan *et al.* (1988) *Science* 239:1534-1536; and U.K. Patent Publication No. GB 2,276,169, published 21 September 1994); and, any functional fragments obtained from such molecules, wherein such fragments retain immunological binding properties of the parent antibody molecule. The invention further includes antibodies obtained through non-conventional processes, such as phage display.

Preferably, the GBS specific antibodies of the invention are monoclonal antibodies. Monoclonal antibodies of the invention include an antibody composition having a homogeneous antibody population. Monoclonal antibodies of the invention may be obtained from murine hybridomas, as well as human monoclonal antibodies obtained using human rather than murine hybridomas. See, *e.g.*, Cote, *et al.* *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, 1985, p 77.

The antibodies of the invention may be used in diagnostic applications, for example, to detect the presence or absence of GBS in a biological sample. The antibodies of the invention may also be used in the prophylactic or therapeutic treatment of GBS infection.

Antibodies: Gram positive bacteria AI sequences

The Gram positive bacteria AI proteins of the invention may also be used to prepare antibodies specific to the Gram positive bacteria AI proteins. The antibodies are preferably specific to the an oligomeric or hyper-oligomeric form of an AI protein. The invention also includes combinations of antibodies specific to Gram positive bacteria AI proteins selected to provide protection against an increased range of Gram positive bacteria genus, species, serotypes and strain isolates.

For example, a combination may comprise a first and second antibody, wherein said first antibody is specific to a first Gram positive bacteria AI protein and said second antibody is specific to a second Gram positive bacteria AI protein. Preferably, the nucleic acid sequence encoding said first Gram positive bacteria AI protein is not present in a Gram positive bacterial genome comprising a polynucleotide sequence encoding for said second Gram positive bacteria AI protein. Preferably, the nucleic acid sequence encoding said first and second Gram positive bacteria AI proteins are present in the genomes of multiple Gram positive bacteria genus, species, serotypes or strain isolates.

As an example of an instance where the combination of antibodies provides protection against an increased range of bacteria serotypes, the first antibody may be specific to a first GAS AI protein and the second antibody may be specific to a second GAS AI protein. The first GAS AI protein may comprise a GAS AI-1 surface protein, while the second GAS AI protein may comprise a GAS AI-2 or AI-3 surface protein.

As an example of an instance where the combination of antibodies provides protection against an increased range of bacterial species, the first antibody may be specific to a GBS AI protein and the second antibody may be specific to a GAS AI protein. Alternatively, the first antibody may be specific to a GAS AI protein and the second antibody may be specific to a *S. pneumoniae* AI protein.

The Gram positive specific antibodies of the invention include one or more biological moieties that, through chemical or physical means, can bind to or associate with an epitope of a Gram positive bacteria AI polypeptide. The antibodies of the invention include antibodies which specifically bind to a Gram positive bacteria AI protein. The invention includes antibodies obtained from both polyclonal and monoclonal preparations, as well as the following: hybrid (chimeric) antibody molecules (see, for example, Winter *et al.* (1991) *Nature* 349: 293-299; and US Patent No. 4,816,567; F(ab')₂ and F(ab) fragments; F_v molecules (non-covalent heterodimers, see, for example, Inbar *et al.* (1972) *Proc Natl Acad Sci USA* 69:2659-2662; and Ehrlich *et al.* (1980) *Biochem* 19:4091-4096); single-chain F_v molecules (sFv) (see, for example, Huston *et al.* (1988) *Proc Natl Acad Sci USA* 85:5897-5883); dimeric and trimeric antibody fragment constructs; minibodies (see, e.g., Pack *et al.* (1992) *Biochem* 31:1579-1584; Cumber *et al.* (1992) *J Immunology* 149B: 120-126); humanized antibody molecules (see, for example, Riechmann *et al.* (1988) *Nature* 332:323-327; Verhoeven *et al.* (1988) *Science* 239:1534-1536; and U.K. Patent Publication No. GB 2,276,169, published 21 September 1994); and, any functional fragments obtained from such molecules, wherein such fragments retain immunological binding properties of the parent antibody molecule. The invention further includes antibodies obtained through non-conventional processes, such as phage display.

Preferably, the Gram positive specific antibodies of the invention are monoclonal antibodies. Monoclonal antibodies of the invention include an antibody composition having a homogeneous antibody population. Monoclonal antibodies of the invention may be obtained from murine hybridomas, as well as human monoclonal antibodies obtained using human rather than murine

The antibodies of the invention may be used in diagnostic applications, for example, to detect the presence or absence of Gram positive bacteria in a biological sample. The antibodies of the invention may also be used in the prophylactic or therapeutic treatment of Gram positive bacteria infection.

Nucleic Acids

The invention provides nucleic acids encoding the Gram positive bacteria sequences and/or the hybrid fusion polypeptides of the invention. The invention also provides nucleic acid encoding the GBS antigens and/or the hybrid fusion polypeptides of the invention. Furthermore, the invention provides nucleic acid which can hybridise to these nucleic acids, preferably under "high stringency" conditions (e.g. 65°C in a 0.1xSSC, 0.5% SDS solution).

Polypeptides of the invention can be prepared by various means (e.g. recombinant expression, purification from cell culture, chemical synthesis, *etc.*) and in various forms (e.g. native, fusions, non-glycosylated, lipidated, *etc.*). They are preferably prepared in substantially pure form (*i.e.* substantially free from other GAS or host cell proteins).

Nucleic acid according to the invention can be prepared in many ways (e.g. by chemical synthesis, from genomic or cDNA libraries, from the organism itself, *etc.*) and can take various forms (e.g. single stranded, double stranded, vectors, probes, *etc.*). They are preferably prepared in substantially pure form (*i.e.* substantially free from other GBS or host cell nucleic acids).

The term "nucleic acid" includes DNA and RNA, and also their analogues, such as those containing modified backbones (e.g. phosphorothioates, *etc.*), and also peptide nucleic acids (PNA), *etc.* The invention includes nucleic acid comprising sequences complementary to those described above (e.g. for antisense or probing purposes).

The invention also provides a process for producing a polypeptide of the invention, comprising the step of culturing a host cell transformed with nucleic acid of the invention under conditions which induce polypeptide expression.

The invention provides a process for producing a polypeptide of the invention, comprising the step of synthesising at least part of the polypeptide by chemical means.

The invention provides a process for producing nucleic acid of the invention, comprising the step of amplifying nucleic acid using a primer-based amplification method (e.g. PCR).

The invention provides a process for producing nucleic acid of the invention, comprising the step of synthesising at least part of the nucleic acid by chemical means.

Purification and Recombinant Expression

The Gram positive bacteria AI proteins of the invention may be isolated from the native Gram positive bacteria, or they may be recombinantly produced, for instance in a heterologous host. For example, the GAS, GBS, and *S. pneumoniae* antigens of the invention may be isolated from

~~*Streptococcus agalactiae*, *S. pyogenes*, *S. pneumoniae*~~, or they may be recombinantly produced, for instance, in a heterologous host. Preferably, the GBS antigens are prepared using a heterologous host.

The heterologous host may be prokaryotic (e.g. a bacterium) or eukaryotic. It is preferably *E.coli*, but other suitable hosts include *Bacillus subtilis*, *Vibrio cholerae*, *Salmonella typhi*, *Salmonella typhimurium*, *Neisseria lactamica*, *Neisseria cinerea*, *Mycobacteria* (e.g. *M.tuberculosis*), *S. gordonii*, *L. lactis*, yeasts, etc.

Recombinant production of polypeptides is facilitated by adding a tag protein to the Gram positive bacteria AI sequence to be expressed as a fusion protein comprising the tag protein and the Gram positive bacteria antigen. For example, recombinant production of polypeptides is facilitated by adding a tag protein to the GBS antigen to be expressed as a fusion protein comprising the tag protein and the GBS antigen. Such tag proteins can facilitate purification, detection and stability of the expressed protein. Tag proteins suitable for use in the invention include a polyarginine tag (Arg-tag), polyhistidine tag (His-tag), FLAG-tag, Strep-tag, c-myc-tag, S-tag, calmodulin-binding peptide, cellulose-binding domain, SBP-tag,, chitin-binding domain, glutathione S-transferase-tag (GST), maltose-binding protein, transcription termination anti-terminiation factor (NusA), *E. coli* thioredoxin (TrxA) and protein disulfide isomerase I (DsbA). Preferred tag proteins include His-tag and GST. A full discussion on the use of tag proteins can be found at Terpe et al., "Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems", Appl Microbiol Biotechnol (2003) 60:523 – 533.

After purification, the tag proteins may optionally be removed from the expressed fusion protein, i.e., by specifically tailored enzymatic treatments known in the art. Commonly used proteases include enterokinase, tobacco etch virus (TEV), thrombin, and factor X_a.

GBS polysaccharides

The compositions of the invention may be further improved by including GBS polysaccharides. Preferably, the GBS antigen and the saccharide each contribute to the immunological response in a recipient. The combination is particularly advantageous where the saccharide and polypeptide provide protection from different GBS serotypes.

The combined antigens may be present as a simple combination where separate saccharide and polypeptide antigens are administered together, or they may be present as a conjugated combination, where the saccharide and polypeptide antigens are covalently linked to each other.

Thus the invention provides an immunogenic composition comprising (i) one or more GBS AI proteins and (ii) one or more GBS saccharide antigens. The polypeptide and the polysaccharide may advantageously be covalently linked to each other to form a conjugate.

Between them, the combined polypeptide and saccharide antigens preferably cover (or provide protection from) two or more GBS serotypes (e.g. 2, 3, 4, 5, 6, 7, 8 or more serotypes). The serotypes of the polypeptide and saccharide antigens may or may not overlap. For example, the polypeptide might protect against serogroup II or V, while the saccharide protects against either serogroups Ia, Ib, or III. Preferred combinations protect against the following groups of serotypes:

(1) serotypes Ia and Ib, (2) serotypes Ia and II, (3) serotypes Ia and III, (4) serotypes Ia and IV, (5) serotypes Ia and V, (6) serotypes Ia and VI, (7) serotypes Ia and VII, (8) serotypes Ia and VIII, (9) serotypes Ib and II, (10) serotypes Ib and III, (11) serotypes Ib and IV, (12) serotypes Ib and V, (13) serotypes Ib and VI, (14) serotypes Ib and VII, (15) serotypes Ib and VIII, (16) serotypes II and III, (17) serotypes II and IV, (18) serotypes II and V, (19) serotypes II and VI, (20) serotypes II and VII, (21) serotypes II and VIII, (22) serotypes III and IV, (23) serotypes III and V, (24) serotypes III and VI, (25) serotypes III and VII, (26) serotypes III and VIII, (27) serotypes IV and V, (28) serotypes IV and VI, (29) serotypes IV and VII, (30) serotypes IV and VIII, (31) serotypes V and VI, (32) serotypes V and VII, (33) serotypes V and VIII, (34) serotypes VI and VII, (35) serotypes VI and VIII, and (36) serotypes VII and VIII.

Still more preferably, the combinations protect against the following groups of serotypes: (1) serotypes Ia and II, (2) serotypes Ia and V, (3) serotypes Ib and II, (4) serotypes Ib and V, (5) serotypes III and II, and (6) serotypes III and V. Most preferably, the combinations protect against serotypes III and V.

Protection against serotypes II and V is preferably provided by polypeptide antigens. Protection against serotypes Ia, Ib and/or III may be polypeptide or saccharide antigens.

Immunogenic compositions and medicaments

Compositions of the invention are preferably immunogenic compositions, and are more preferably vaccine compositions. The pH of the composition is preferably between 6 and 8, preferably about 7. The pH may be maintained by the use of a buffer. The composition may be sterile and/or pyrogen-free. The composition may be isotonic with respect to humans.

Vaccines according to the invention may either be prophylactic (*i.e.* to prevent infection) or therapeutic (*i.e.* to treat infection), but will typically be prophylactic. Accordingly, the invention includes a method for the therapeutic or prophylactic treatment of a Gram positive bacteria infection in an animal susceptible to such gram positive bacterial infection comprising administering to said animal a therapeutic or prophylactic amount of the immunogenic composition of the invention. For example, the invention includes a method for the therapeutic or prophylactic treatment of a *Streptococcus agalactiae*, *S. pyogenes*, or *S. pneumoniae* infection in an animal susceptible to streptococcal infection comprising administering to said animal a therapeutic or prophylactic amount of the immunogenic compositions of the invention.

The invention also provides a composition of the invention for use of the compositions described herein as a medicament. The medicament is preferably able to raise an immune response in a mammal (*i.e.* it is an immunogenic composition) and is more preferably a vaccine.

The invention also provides the use of the compositions of the invention in the manufacture of a medicament for raising an immune response in a mammal. The medicament is preferably a vaccine.

The invention also provides kits comprising one or more containers of compositions of the invention. Compositions can be in liquid form or can be lyophilized, as can individual antigens. Suitable containers for the compositions include, for example, bottles, vials, syringes, and test tubes.

Containers can be formed from a variety of materials, including glass or plastic. A container may have a sterile access port (for example, the container may be an intravenous solution bag or a vial having a stopper pierceable by a hypodermic injection needle). The composition may comprise a first component comprising one or more Gram positive bacteria AI proteins. Preferably, the AI proteins are surface AI proteins. Preferably, the AI surface proteins are in an oligomeric or hyperoligomeric form. For example, the first component comprises a combination of GBS antigens or GAS antigens, or *S. pneumoniae* antigens. Preferably said combination includes GBS 80. Preferably GBS 80 is present in an oligomeric or hyperoligomeric form.

The kit can further comprise a second container comprising a pharmaceutically-acceptable buffer, such as phosphate-buffered saline, Ringer's solution, or dextrose solution. It can also contain other materials useful to the end-user, including other buffers, diluents, filters, needles, and syringes. The kit can also comprise a second or third container with another active agent, for example an antibiotic.

The kit can also comprise a package insert containing written instructions for methods of inducing immunity against *S. agalactiae* and/or *S. pyogenes* and/or *S. pneumoniae* or for treating *S. agalactiae* and/or *S. pyogenes* and/or *S. pneumoniae* infections. The package insert can be an unapproved draft package insert or can be a package insert approved by the Food and Drug Administration (FDA) or other regulatory body.

The invention also provides a delivery device pre-filled with the immunogenic compositions of the invention.

The invention also provides a method for raising an immune response in a mammal comprising the step of administering an effective amount of a composition of the invention. The immune response is preferably protective and preferably involves antibodies and/or cell-mediated immunity. This immune response will preferably induce long lasting (e.g., neutralising) antibodies and a cell mediated immunity that can quickly respond upon exposure to one or more GBS and/or GAS and/or *S. pneumoniae* antigens. The method may raise a booster response.

The invention provides a method of neutralizing GBS, GAS, or *S. pneumoniae* infection in a mammal comprising the step of administering to the mammal an effective amount of the immunogenic compositions of the invention, a vaccine of the invention, or antibodies which recognize an immunogenic composition of the invention.

The mammal is preferably a human. Where the vaccine is for prophylactic use, the human is preferably a female (either of child bearing age or a teenager). Alternatively, the human may be elderly (e.g., over the age of 50, 55, 60, 65, 70 or 75) and may have an underlying disease such as diabetes or cancer. Where the vaccine is for therapeutic use, the human is preferably a pregnant female or an elderly adult.

These uses and methods are preferably for the prevention and/or treatment of a disease caused by *Streptococcus agalactiae*, or *S. pyogenes*, or *S. pneumoniae*. The compositions may also be

effective against other streptococcal bacteria. The compositions may also be effective against other Gram positive bacteria.

One way of checking efficacy of therapeutic treatment involves monitoring Gram positive bacterial infection after administration of the composition of the invention. One way of checking efficacy of prophylactic treatment involves monitoring immune responses against the Gram positive bacterial antigens in the compositions of the invention after administration of the composition.

One way of checking efficacy of therapeutic treatment involves monitoring GBS infection after administration of the composition of the invention. One way of checking efficacy of prophylactic treatment involves monitoring immune responses against the GBS antigens in the compositions of the invention after administration of the composition.

A way of assessing the immunogenicity of the component proteins of the immunogenic compositions of the present invention is to express the proteins recombinantly and to screen patient sera or mucosal secretions by immunoblot. A positive reaction between the protein and the patient serum indicates that the patient has previously mounted an immune response to the protein in question- that is, the protein is an immunogen. This method may also be used to identify immunodominant proteins and/or epitopes.

Another way of checking efficacy of therapeutic treatment involves monitoring GBS or GAS or *S pneumoniae* infection after administration of the compositions of the invention. One way of checking efficacy of prophylactic treatment involves monitoring immune responses both systemically (such as monitoring the level of IgG1 and IgG2a production) and mucosally (such as monitoring the level of IgA production) against the GBS and/or GAS and/or *S pneumoniae* antigens in the compositions of the invention after administration of the composition. Typically, GBS and/or GAS and/or *S pneumoniae* serum specific antibody responses are determined post-immunization but pre-challenge whereas mucosal GBS and/or GAS and/or *S pneumoniae* specific antibody body responses are determined post-immunization and post-challenge.

The vaccine compositions of the present invention can be evaluated in *in vitro* and *in vivo* animal models prior to host, *e.g.*, human, administration.

The efficacy of immunogenic compositions of the invention can also be determined *in vivo* by challenging animal models of GBS and/or GAS and/or *S pneumoniae* infection, *e.g.*, guinea pigs or mice, with the immunogenic compositions. The immunogenic compositions may or may not be derived from the same serotypes as the challenge serotypes. Preferably the immunogenic compositions are derivable from the same serotypes as the challenge serotypes. More preferably, the immunogenic composition and/or the challenge serotypes are derivable from the group of GBS and/or GAS and/or *S pneumoniae* serotypes.

In vivo efficacy models include but are not limited to: (i) A murine infection model using human GBS and/or GAS and/or *S pneumoniae* serotypes; (ii) a murine disease model which is a murine model using a mouse-adapted GBS and/or GAS and/or *S pneumoniae* strain, such as those

strains outlined above which is particularly virulent in mice and (iii) a primate model using human GBS or GAS or S pneumoniae isolates.

The immune response may be one or both of a TH1 immune response and a TH2 response.

The immune response may be an improved or an enhanced or an altered immune response.

5 The immune response may be one or both of a systemic and a mucosal immune response.

Preferably the immune response is an enhanced system and/or mucosal response.

An enhanced systemic and/or mucosal immunity is reflected in an enhanced TH1 and/or TH2 immune response. Preferably, the enhanced immune response includes an increase in the production of IgG1 and/or IgG2a and/or IgA

10 Preferably the mucosal immune response is a TH2 immune response. Preferably, the mucosal immune response includes an increase in the production of IgA.

Activated TH2 cells enhance antibody production and are therefore of value in responding to extracellular infections. Activated TH2 cells may secrete one or more of IL-4, IL-5, IL-6, and IL-10.

15 A TH2 immune response may result in the production of IgG1, IgE, IgA and memory B cells for future protection.

A TH2 immune response may include one or more of an increase in one or more of the cytokines associated with a TH2 immune response (such as IL-4, IL-5, IL-6 and IL-10), or an increase in the production of IgG1, IgE, IgA and memory B cells. Preferably, the enhanced TH2 immune response will include an increase in IgG1 production.

20 A TH1 immune response may include one or more of an increase in CTLs, an increase in one or more of the cytokines associated with a TH1 immune response (such as IL-2, IFN γ , and TNF β), an increase in activated macrophages, an increase in NK activity, or an increase in the production of IgG2a. Preferably, the enhanced TH1 immune response will include an increase in IgG2a production.

25 Immunogenic compositions of the invention, in particular, immunogenic composition comprising one or more GAS antigens of the present invention may be used either alone or in combination with other GAS antigens optionally with an immunoregulatory agent capable of eliciting a Th1 and/or Th2 response.

30 Compositions of the invention will generally be administered directly to a patient. Certain routes may be favored for certain compositions, as resulting in the generation of a more effective immune response, preferably a CMI response, or as being less likely to induce side effects, or as being easier for administration. Direct delivery may be accomplished by parenteral injection (e.g. subcutaneously, intraperitoneally, intradermally, intravenously, intramuscularly, or to the interstitial space of a tissue), or by rectal, oral (e.g. tablet, spray), vaginal, topical, transdermal (e.g. see WO 99/27961) or transcutaneous (e.g. see WO 02/074244 and WO 02/064162), intranasal (e.g. see 35 WO03/028760), ocular, aural, pulmonary or other mucosal administration.

The invention may be used to elicit systemic and/or mucosal immunity.

In one particularly preferred embodiment, the immunogenic composition comprises one or more GBS or GAS or S pneumoniae antigen(s) which elicits a neutralising antibody response and one or more GBS or GAS or S pneumoniae antigen(s) which elicit a cell mediated immune response. In this way, the neutralising antibody response prevents or inhibits an initial GBS or GAS or S pneumoniae infection while the cell-mediated immune response capable of eliciting an enhanced Th1 cellular response prevents further spreading of the GBS or GAS or S pneumoniae infection. Preferably, the immunogenic composition comprises one or more GBS or GAS or S pneumoniae surface antigens and one or more GBS or GAS or S pneumoniae cytoplasmic antigens. Preferably the immunogenic composition comprises one or more GBS or GAS or S pneumoniae surface antigens or the like and one or other antigens, such as a cytoplasmic antigen capable of eliciting a Th1 cellular response.

Dosage treatment can be a single dose schedule or a multiple dose schedule. Multiple doses may be used in a primary immunisation schedule and/or in a booster immunisation schedule. In a multiple dose schedule the various doses may be given by the same or different routes *e.g.* a parenteral prime and mucosal boost, a mucosal prime and parenteral boost, *etc.*

The compositions of the invention may be prepared in various forms. For example, the compositions may be prepared as injectables, either as liquid solutions or suspensions. Solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection can also be prepared (*e.g.* a lyophilised composition). The composition may be prepared for topical administration *e.g.* as an ointment, cream or powder. The composition may be prepared for oral administration *e.g.* as a tablet or capsule, as a spray, or as a syrup (optionally flavoured). The composition may be prepared for pulmonary administration *e.g.* as an inhaler, using a fine powder or a spray. The composition may be prepared as a suppository or pessary. The composition may be prepared for nasal, aural or ocular administration *e.g.* as drops. The composition may be in kit form, designed such that a combined composition is reconstituted just prior to administration to a patient. Such kits may comprise one or more antigens in liquid form and one or more lyophilised antigens.

Immunogenic compositions used as vaccines comprise an immunologically effective amount of antigen(s), as well as any other components, such as antibiotics, as needed. By 'immunologically effective amount', it is meant that the administration of that amount to an individual, either in a single dose or as part of a series, is effective for treatment or prevention, or increases a measurable immune response or prevents or reduces a clinical symptom. This amount varies depending upon the health and physical condition of the individual to be treated, age, the taxonomic group of individual to be treated (*e.g.* non-human primate, primate, *etc.*), the capacity of the individual's immune system to synthesise antibodies, the degree of protection desired, the formulation of the vaccine, the treating doctor's assessment of the medical situation, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials.

Further Components of the Composition

The composition of the invention will typically, in addition to the components mentioned above, comprise one or more 'pharmaceutically acceptable carriers', which include any carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolised macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, and lipid aggregates (such as oil droplets or liposomes). Such carriers are well known to those of ordinary skill in the art. The vaccines may also contain diluents, such as water, saline, glycerol, *etc.* Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present. A thorough discussion of pharmaceutically acceptable excipients is available in Gennaro (2000) *Remington: The Science and Practice of Pharmacy*. 20th ed., ISBN: 0683306472.

Adjuvants

Vaccines of the invention may be administered in conjunction with other immunoregulatory agents. In particular, compositions will usually include an adjuvant. Adjuvants for use with the invention include, but are not limited to, one or more of the following set forth below:

A. Mineral Containing Compositions

Mineral containing compositions suitable for use as adjuvants in the invention include mineral salts, such as aluminum salts and calcium salts. The invention includes mineral salts such as hydroxides (*e.g.* oxyhydroxides), phosphates (*e.g.* hydroxyphosphates, orthophosphates), sulfates, *etc.* (*e.g.* see chapters 8 & 9 of *Vaccine Design...* (1995) eds. Powell & Newman. ISBN: 030644867X. Plenum.), or mixtures of different mineral compounds (*e.g.* a mixture of a phosphate and a hydroxide adjuvant, optionally with an excess of the phosphate), with the compounds taking any suitable form (*e.g.* gel, crystalline, amorphous, *etc.*), and with adsorption to the salt(s) being preferred. The mineral containing compositions may also be formulated as a particle of metal salt (WO 00/23105).

Aluminum salts may be included in vaccines of the invention such that the dose of Al^{3+} is between 0.2 and 1.0 mg per dose.

B. Oil-Emulsions

Oil-emulsion compositions suitable for use as adjuvants in the invention include squalene-water emulsions, such as MF59 (5% Squalene, 0.5% Tween 80, and 0.5% Span 85, formulated into submicron particles using a microfluidizer). See WO90/14837. See also, Podda, "The adjuvanted influenza vaccines with novel adjuvants: experience with the MF59-adjuvanted vaccine", *Vaccine* (2001) 19: 2673-2680; Frey et al., "Comparison of the safety, tolerability, and immunogenicity of a MF59-adjuvanted influenza vaccine and a non-adjuvanted influenza vaccine in non-elderly adults", *Vaccine* (2003) 21:4234-4237. MF59 is used as the adjuvant in the FLUAD™ influenza virus trivalent subunit vaccine.

Particularly preferred adjuvants for use in the compositions are submicron oil-in-water emulsions. Preferred submicron oil-in-water emulsions for use herein are squalene/water emulsions optionally containing varying amounts of MTP-PE, such as a submicron oil-in-water emulsion containing 4-5% w/v squalene, 0.25-1.0% w/v Tween 80™ (polyoxyethylsorbitan monooleate), and/or 0.25-1.0% Span 85™ (sorbitan trioleate), and, optionally, N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-*sn*-glycero-3-hydroxyphosphoryloxy)-ethylamine (MTP-PE), for example, the submicron oil-in-water emulsion known as "MF59" (International Publication No. WO 90/14837; US Patent Nos. 6,299,884 and 6,451,325, incorporated herein by reference in their entireties; and Ott et al., "MF59 -- Design and Evaluation of a Safe and Potent Adjuvant for Human Vaccines" in *Vaccine Design: The Subunit and Adjuvant Approach* (Powell, M.F. and Newman, M.J. eds.) Plenum Press, New York, 1995, pp. 277-296). MF59 contains 4-5% w/v Squalene (e.g. 4.3%), 0.25-0.5% w/v Tween 80™, and 0.5% w/v Span 85™ and optionally contains various amounts of MTP-PE, formulated into submicron particles using a microfluidizer such as Model 110Y microfluidizer (Microfluidics, Newton, MA). For example, MTP-PE may be present in an amount of about 0-500 µg/dose, more preferably 0-250 µg/dose and most preferably, 0-100 µg/dose. As used herein, the term "MF59-0" refers to the above submicron oil-in-water emulsion lacking MTP-PE, while the term MF59-MTP denotes a formulation that contains MTP-PE. For instance, "MF59-100" contains 100 µg MTP-PE per dose, and so on. MF69, another submicron oil-in-water emulsion for use herein, contains 4.3% w/v squalene, 0.25% w/v Tween 80™, and 0.75% w/v Span 85™ and optionally MTP-PE. Yet another submicron oil-in-water emulsion is MF75, also known as SAF, containing 10% squalene, 0.4% Tween 80™, 5% pluronic-blocked polymer L121, and thr-MDP, also microfluidized into a submicron emulsion. MF75-MTP denotes an MF75 formulation that includes MTP, such as from 100-400 µg MTP-PE per dose.

Submicron oil-in-water emulsions, methods of making the same and immunostimulating agents, such as muramyl peptides, for use in the compositions, are described in detail in International Publication No. WO 90/14837 and US Patent Nos. 6,299,884 and 6,451,325, incorporated herein by reference in their entireties.

Complete Freund's adjuvant (CFA) and incomplete Freund's adjuvant (IFA) may also be used as adjuvants in the invention.

C. Saponin Formulations

Saponin formulations, may also be used as adjuvants in the invention. Saponins are a heterologous group of sterol glycosides and triterpenoid glycosides that are found in the bark, leaves, stems, roots and even flowers of a wide range of plant species. Saponin from the bark of the *Quillaja saponaria* Molina tree have been widely studied as adjuvants. Saponin can also be commercially obtained from *Smilax ornata* (sarsapilla), *Gypsophilla paniculata* (brides veil), and *Saponaria officianalis* (soap root). Saponin adjuvant formulations include purified formulations, such as QS21, as well as lipid formulations, such as ISCOMs.

Saponin compositions have been purified using High Performance Thin Layer Chromatography (HP-LC) and Reversed Phase High Performance Liquid Chromatography (RP-HPLC). Specific purified fractions using these techniques have been identified, including QS7, QS17, QS18, QS21, QH-A, QH-B and QH-C. Preferably, the saponin is QS21. A method of production of QS21 is disclosed in US Patent No. 5,057,540. Saponin formulations may also comprise a sterol, such as cholesterol (see WO96/33739).

Combinations of saponins and cholesterol can be used to form unique particles called Immunostimulating Complexes (ISCOMs). ISCOMs typically also include a phospholipid such as phosphatidylethanolamine or phosphatidylcholine. Any known saponin can be used in ISCOMs. Preferably, the ISCOM includes one or more of Quil A, QHA and QHC. ISCOMs are further described in EP0109942, WO 96/11711 and WO 96/33739. Optionally, the ISCOMS may be devoid of additional detergent. See WO 00/07621.

A review of the development of saponin based adjuvants can be found at Barr, et al., "ISCOMs and other saponin based adjuvants", *Advanced Drug Delivery Reviews* (1998) 32:247-271. See also Sjolander, et al., "Uptake and adjuvant activity of orally delivered saponin and ISCOM vaccines", *Advanced Drug Delivery Reviews* (1998) 32:321-338.

D. *Virosomes and Virus Like Particles (VLPs)*

Virosomes and Virus Like Particles (VLPs) can also be used as adjuvants in the invention. These structures generally contain one or more proteins from a virus optionally combined or formulated with a phospholipid. They are generally non-pathogenic, non-replicating and generally do not contain any of the native viral genome. The viral proteins may be recombinantly produced or isolated from whole viruses. These viral proteins suitable for use in virosomes or VLPs include proteins derived from influenza virus (such as HA or NA), Hepatitis B virus (such as core or capsid proteins), Hepatitis E virus, measles virus, Sindbis virus, Rotavirus, Foot-and-Mouth Disease virus, Retrovirus, Norwalk virus, human Papilloma virus, HIV, RNA-phages, Q β -phage (such as coat proteins), GA-phage, fr-phage, AP205 phage, and Ty (such as retrotransposon Ty protein p1). VLPs are discussed further in WO 03/024480, WO 03/024481, and Niikura et al., "Chimeric Recombinant Hepatitis E Virus-Like Particles as an Oral Vaccine Vehicle Presenting Foreign Epitopes", *Virology* (2002) 293:273-280; Lenz et al., "Papillomavirus-Like Particles Induce Acute Activation of Dendritic Cells", *Journal of Immunology* (2001) 5246-5355; Pinto, et al., "Cellular Immune Responses to Human Papillomavirus (HPV)-16 L1 Healthy Volunteers Immunized with Recombinant HPV-16 L1 Virus-Like Particles", *Journal of Infectious Diseases* (2003) 188:327-338; and Gerber et al., "Human Papillomavirus Virus-Like Particles Are Efficient Oral Immunogens when Coadministered with Escherichia coli Heat-Labile Enterotoxin Mutant R192G or CpG", *Journal of Virology* (2001) 75(10):4752-4760. Virosomes are discussed further in, for example, Gluck et al., "New Technology Platforms in the Development of Vaccines for the Future", *Vaccine* (2002) 20:B10-B16. Immunopotentiating reconstituted influenza virosomes (IRIV) are used as the subunit antigen

E. Bacterial or Microbial Derivatives

Adjuvants suitable for use in the invention include bacterial or microbial derivatives such as:

(1) Non-toxic derivatives of enterobacterial lipopolysaccharide (LPS)

Such derivatives include Monophosphoryl lipid A (MPL) and 3-O-deacylated MPL (3dMPL).

3dMPL is a mixture of 3 De-O-acylated monophosphoryl lipid A with 4, 5 or 6 acylated chains. A preferred "small particle" form of 3 De-O-acylated monophosphoryl lipid A is disclosed in EP 0 689 454. Such "small particles" of 3dMPL are small enough to be sterile filtered through a 0.22 micron membrane (see EP 0 689 454). Other non-toxic LPS derivatives include monophosphoryl lipid A mimics, such as aminoalkyl glucosaminide phosphate derivatives e.g. RC-529. See Johnson *et al.* (1999) *Bioorg Med Chem Lett* 9:2273-2278.

(2) Lipid A Derivatives

Lipid A derivatives include derivatives of lipid A from *Escherichia coli* such as OM-174.

OM-174 is described for example in Meraldi *et al.*, "OM-174, a New Adjuvant with a Potential for Human Use, Induces a Protective Response with Administered with the Synthetic C-Terminal Fragment 242-310 from the circumsporozoite protein of *Plasmodium berghei*", *Vaccine* (2003) 21:2485-2491; and Pajak, *et al.*, "The Adjuvant OM-174 induces both the migration and maturation of murine dendritic cells in vivo", *Vaccine* (2003) 21:836-842.

(3) Immunostimulatory oligonucleotides

Immunostimulatory oligonucleotides suitable for use as adjuvants in the invention include nucleotide sequences containing a CpG motif (a sequence containing an unmethylated cytosine followed by guanosine and linked by a phosphate bond). Bacterial double stranded RNA or oligonucleotides containing palindromic or poly(dG) sequences have also been shown to be immunostimulatory.

The CpG's can include nucleotide modifications/analogs such as phosphorothioate modifications and can be double-stranded or single-stranded. Optionally, the guanosine may be replaced with an analog such as 2'-deoxy-7-deazaguanosine. See Kandimalla, *et al.*, "Divergent synthetic nucleotide motif recognition pattern: design and development of potent immunomodulatory oligodeoxyribonucleotide agents with distinct cytokine induction profiles", *Nucleic Acids Research* (2003) 31(9): 2393-2400; WO02/26757 and WO99/62923 for examples of possible analog substitutions. The adjuvant effect of CpG oligonucleotides is further discussed in Krieg, "CpG motifs: the active ingredient in bacterial extracts?", *Nature Medicine* (2003) 9(7): 831-835; McCluskie, *et al.*, "Parenteral and mucosal prime-boost immunization strategies in mice with hepatitis B surface antigen and CpG DNA", *FEMS Immunology and Medical Microbiology* (2002) 32:179-185; WO98/40100; US Patent No. 6,207,646; US Patent No. 6,239,116 and US Patent No. 6,429,199.

The CpG sequence may be directed to TLR9, such as the motif GTCGTT or TTCGTT. See Kandimalla, *et al.*, "Toll-like receptor 9: modulation of recognition and cytokine induction by novel

synthetic CpG DNAs” Biochemical Society Transactions (2003) 31 (part 3): 654-658. The CpG sequence may be specific for inducing a Th1 immune response, such as a CpG-A ODN, or it may be more specific for inducing a B cell response, such as a CpG-B ODN. CpG-A and CpG-B ODNs are discussed in Blackwell, et al., “CpG-A-Induced Monocyte IFN-gamma-Inducible Protein-10 Production is Regulated by Plasmacytoid Dendritic Cell Derived IFN-alpha”, J. Immunol. (2003) 170(8):4061-4068; Krieg, “From A to Z on CpG”, TRENDS in Immunology (2002) 23(2): 64-65 and WO01/95935. Preferably, the CpG is a CpG-A ODN.

Preferably, the CpG oligonucleotide is constructed so that the 5' end is accessible for receptor recognition. Optionally, two CpG oligonucleotide sequences may be attached at their 3' ends to form “immunomers”. See, for example, Kandimalla, et al., “Secondary structures in CpG oligonucleotides affect immunostimulatory activity”, BBRC (2003) 306:948-953; Kandimalla, et al., “Toll-like receptor 9: modulation of recognition and cytokine induction by novel synthetic CpG DNAs”, Biochemical Society Transactions (2003) 31(part 3):664-658; Bhagat et al., “CpG penta- and hexadeoxyribonucleotides as potent immunomodulatory agents” BBRC (2003) 300:853-861 and WO 03/035836.

(4) *ADP-ribosylating toxins and detoxified derivatives thereof.*

Bacterial ADP-ribosylating toxins and detoxified derivatives thereof may be used as adjuvants in the invention. Preferably, the protein is derived from *E. coli* (i.e., *E. coli* heat labile enterotoxin “LT”, cholera (“CT”), or pertussis (“PT”). The use of detoxified ADP-ribosylating toxins as mucosal adjuvants is described in WO95/17211 and as parenteral adjuvants in WO98/42375. Preferably, the adjuvant is a detoxified LT mutant such as LT-K63, LT-R72, and LTR192G. The use of ADP-ribosylating toxins and detoxified derivatives thereof, particularly LT-K63 and LT-R72, as adjuvants can be found in the following references, each of which is specifically incorporated by reference herein in their entirety: Beignon, et al., “The LTR72 Mutant of Heat-Labile Enterotoxin of *Escherichia coli* Enhances the Ability of Peptide Antigens to Elicit CD4+ T Cells and Secrete Gamma Interferon after Coapplication onto Bare Skin”, Infection and Immunity (2002) 70(6):3012-3019; Pizza, et al., “Mucosal vaccines: non toxic derivatives of LT and CT as mucosal adjuvants”, Vaccine (2001) 19:2534-2541; Pizza, et al., “LTK63 and LTR72, two mucosal adjuvants ready for clinical trials” Int. J. Med. Microbiol (2000) 290(4-5):455-461; Scharton-Kersten et al., “Transcutaneous Immunization with Bacterial ADP-Ribosylating Exotoxins, Subunits and Unrelated Adjuvants”, Infection and Immunity (2000) 68(9):5306-5313; Ryan et al., “Mutants of *Escherichia coli* Heat-Labile Toxin Act as Effective Mucosal Adjuvants for Nasal Delivery of an Acellular Pertussis Vaccine: Differential Effects of the Nontoxic AB Complex and Enzyme Activity on Th1 and Th2 Cells” Infection and Immunity (1999) 67(12):6270-6280; Partidos et al., “Heat-labile enterotoxin of *Escherichia coli* and its site-directed mutant LTK63 enhance the proliferative and cytotoxic T-cell responses to intranasally co-immunized synthetic peptides”, Immunol. Lett. (1999) 67(3):209-216; Peppoloni et al., “Mutants of the *Escherichia coli* heat-labile enterotoxin as safe and strong adjuvants for intranasal delivery of vaccines”, Vaccines (2003) 2(2):285-293; and Pine et al., (2002) “Intranasal

-241-

immunization with influenza vaccine and a detoxified mutant of heat labile enterotoxin from *Escherichia coli* (LTK63)" J. Control Release (2002) 85(1-3):263-270. Numerical reference for amino acid substitutions is preferably based on the alignments of the A and B subunits of ADP-ribosylating toxins set forth in Domenighini et al., Mol. Microbiol (1995) 15(6):1165-1167, specifically incorporated herein by reference in its entirety.

F. Bioadhesives and Mucoadhesives

Bioadhesives and mucoadhesives may also be used as adjuvants in the invention. Suitable bioadhesives include esterified hyaluronic acid microspheres (Singh et al. (2001) J. Cont. Rele. 70:267-276) or mucoadhesives such as cross-linked derivatives of poly(acrylic acid), polyvinyl alcohol, polyvinyl pyrrolidone, polysaccharides and carboxymethylcellulose. Chitosan and derivatives thereof may also be used as adjuvants in the invention. E.g. WO99/27960.

G. Microparticles

Microparticles may also be used as adjuvants in the invention. Microparticles (i.e. a particle of ~100nm to ~150µm in diameter, more preferably ~200nm to ~30µm in diameter, and most preferably ~500nm to ~10µm in diameter) formed from materials that are biodegradable and non-toxic (e.g. a poly(α-hydroxy acid), a polyhydroxybutyric acid, a polyorthoester, a polyanhydride, a polycaprolactone, etc.), with poly(lactide-co-glycolide) are preferred, optionally treated to have a negatively-charged surface (e.g. with SDS) or a positively-charged surface (e.g. with a cationic detergent, such as CTAB).

H. Liposomes

Examples of liposome formulations suitable for use as adjuvants are described in US Patent No. 6,090,406, US Patent No. 5,916,588, and EP 0 626 169.

I. Polyoxyethylene ether and Polyoxyethylene Ester Formulations

Adjuvants suitable for use in the invention include polyoxyethylene ethers and polyoxyethylene esters. WO99/52549. Such formulations further include polyoxyethylene sorbitan ester surfactants in combination with an octoxynol (WO01/21207) as well as polyoxyethylene alkyl ethers or ester surfactants in combination with at least one additional non-ionic surfactant such as an octoxynol (WO 01/21152).

Preferred polyoxyethylene ethers are selected from the following group: polyoxyethylene-9-lauryl ether (laureth 9), polyoxyethylene-9-stearyl ether, polyoxyethylene-8-stearyl ether, polyoxyethylene-4-lauryl ether, polyoxyethylene-35-lauryl ether, and polyoxyethylene-23-lauryl ether.

J. Polyphosphazene (PCPP)

PCPP formulations are described, for example, in Andrianov et al., "Preparation of hydrogel microspheres by coacervation of aqueous polyphosphazene solutions", Biomaterials (1998) 19(1-3):109-115 and Payne et al., "Protein Release from Polyphosphazene Matrices", Adv. Drug. Delivery Review (1998) 31(3):185-196.

pe K. Muramyl peptides

Examples of muramyl peptides suitable for use as adjuvants in the invention include N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-normuramyl-l-alanyl-d-isoglutamine (nor-MDP), and N-acetylmuramyl-l-alanyl-d-isoglutaminyl-l-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3-hydroxyphosphoryloxy)-ethylamine MTP-PE).

L. Imidazoquinolone Compounds.

Examples of imidazoquinolone compounds suitable for use adjuvants in the invention include Imiquamod and its homologues, described further in Stanley, "Imiquimod and the imidazoquinolones: mechanism of action and therapeutic potential" Clin Exp Dermatol (2002) 27(7):571-577 and Jones, "Resiquimod 3M", Curr Opin Investig Drugs (2003) 4(2):214-218.

The invention may also comprise combinations of aspects of one or more of the adjuvants identified above. For example, the following adjuvant compositions may be used in the invention:

- (1) a saponin and an oil-in-water emulsion (WO 99/11241);
- (2) a saponin (e.g., QS21) + a non-toxic LPS derivative (e.g. 3dMPL) (see WO 94/00153);
- (3) a saponin (e.g., QS21) + a non-toxic LPS derivative (e.g. 3dMPL) + a cholesterol;
- (4) a saponin (e.g. QS21) + 3dMPL + IL-12 (optionally + a sterol) (WO 98/57659);
- (5) combinations of 3dMPL with, for example, QS21 and/or oil-in-water emulsions (See European patent applications 0835318, 0735898 and 0761231);
- (6) SAF, containing 10% Squalane, 0.4% Tween 80, 5% pluronic-block polymer L121, and thr-MDP, either microfluidized into a submicron emulsion or vortexed to generate a larger particle size emulsion.
- (7) RibiTTM adjuvant system (RAS), (Ribi Immunochem) containing 2% Squalene, 0.2% Tween 80, and one or more bacterial cell wall components from the group consisting of monophosphorylipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS), preferably MPL + CWS (DetoxTM);
- (8) one or more mineral salts (such as an aluminum salt) + a non-toxic derivative of LPS (such as 3dPML).
- (9) one or more mineral salts (such as an aluminum salt) + an immunostimulatory oligonucleotide (such as a nucleotide sequence including a CpG motif). Combination No. (9) is a preferred adjuvant combination.

M. Human Immunomodulators

Human immunomodulators suitable for use as adjuvants in the invention include cytokines, such as interleukins (e.g. IL-1, IL-2, IL-4, IL-5, IL-6, IL-7, IL-12, etc.), interferons (e.g. interferon- γ), macrophage colony stimulating factor, and tumor necrosis factor.

Aluminum salts and MF59 are preferred adjuvants for use with injectable influenza vaccines. Bacterial toxins and bioadhesives are preferred adjuvants for use with mucosally-delivered vaccines, such as nasal vaccines.

The immunogenic compositions of the present invention may be administered in combination with an antibiotic treatment regime. In one embodiment, the antibiotic is administered prior to administration of the antigen of the invention or the composition comprising the one or more of the antigens of the invention.

5 In another embodiment, the antibiotic is administered subsequent to the administration of the one or more antigens of the invention or the composition comprising the one or more antigens of the invention. Examples of antibiotics suitable for use in the treatment of the Streptococcal infections of the invention include but are not limited to penicillin or a derivative thereof or clindamycin or the like.

10 Further antigens

The compositions of the invention may further comprise one or more additional Gram positive bacterial antigens which are not associated with an AI. Preferably, the Gram positive bacterial antigens that are not associated with an AI can provide protection across more than one serotype or strain isolate. For example, a first non-AI antigen, in which the first non-AI antigen is at least 90% (*i.e.*, at least 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100%) homologous to the amino acid sequence of a second non-AI antigen, wherein the first and the second non-AI antigen are derived from the genomes of different serotypes of a Gram positive bacteria, may be further included in the compositions. The first non-AI antigen may also be homologous to the amino acid sequence of a third non-AI antigen, such that the first non-AI antigen, the second non-AI antigen, and the third non-AI antigen are derived from the genomes of different serotypes of a Gram positive bacteria. The first non-AI antigen may also be homologous to the amino acid sequence of a fourth non-AI antigen, such that the first non-AI antigen, the second non-AI antigen, the third non-AI antigen, and the fourth non-AI antigen are derived from the genomes of different serotypes of a Gram positive bacteria.

The first non-AI antigen may be GBS 322. The amino acid sequence of GBS 322 across GBS strains from serotypes Ia, Ib, II, III, V, and VIII is greater than 90%. Alternatively, the first non-AI antigen may be GBS 276. The amino acid sequence of GBS 276 across GBS strain from serotypes Ia, Ib, II, III, V, and VIII is greater than 90%. Table 13 provides the percent amino acid sequence identity of GBS 322 and GBS 276 across different GBS strains and serotypes.

Table 13: Conservation of GBS 322 and GBS 276 amino acid sequences

Serotype	Strains	GBS 322		GBS 276	
		cGH	%AA identity	cGH	%AA identity
Ia	090	+	98.60	+	97.90
	A909	+	98.30	+	97.90
	515	+	98.80	+	97.50
	DK1	+		+	
	DK8	+		+	
	Davis	+		+	
Ib	7357b	+		+	
	H36B	+	98.30	+	97.80
II	18RS21	+	100.00	+	99.90
	DK21	+		+	

Serotype	Strains	GBS 322		GBS 276	
		cGH	%AA identity	cGH	%AA identity
III	NEM316	+	100.00	+	97.00
	COH31	+		+	
	D136	+		+	
	M732	+	98.00	+	100.00
	COH1	+	98.30	+	100.00
	M781	+	98.30	+	99.60
No type	CJB110	+	98.60	+	97.90
	1169NT	+	97.40	+	97.90
V	CJB111	+	100.00	+	
	2603	+	100.00	+	100.00
VIII	JM130013	+	100.00	+	97.90
	SMU014	+		+	
total		22/22	98.28+/-0.4	22/22	98.44 +/-1.094

As an example, inclusion of a non-AI protein, GBS 322, in combination with AI antigens GBS 67, GBS 80, and GBS 104 provided protection to newborn mice in an active maternal immunization assay.

5

Table 14: Active maternal immunization assay for a combination of fragments from GBS 322, GBS 80, GBS 104, and GBS 67

GBS strains	Type	FACS (A Mean)			MIX=322+80+104+67		PBS	
		GBS 80	GBS 67	GBS 322	alive/treated	% protection	alive/treated	% protection
515	Ia	0	409	227	39/40	97	6/40	15
7357b-	Ib	91	316	102	19/30	63	1/30	3
DK21	II	0	331	416	25/34	73	17/48	35
5401	II	170	618	135	35/40	87	3/37	8
3050	II	43	460	188	48/48	100	1/30	3
COH1	III	305	0	130	36/36	100	7/40	17
M781	III	65	0	224	30/40	75	4/39	10
2603	V	125	105	313	27/33	82	10/35	28
CJB111	V	370	481	63	25/28	89	4/46	9
JM9130013	VIII	597	83	143	37/39	95	5/40	12
JMU071	VIII	556	79	170	44/50	88	18/50	36
NT1169	NT	0	443	213	12/32	37	11/35	31

In fact, the non-AI GBS 322 antigen may itself provide protection to newborn mice in an active maternal immunization assay.

PCT/US2005/027239

Table 16: Active maternal immunization assay for each of GBS 80 and GBS 322 antigens

GBS strains	Type	GBS 80			GBS 322		
		FACS	Protection (% survival)		FACS	Protection (% survival)	
		Δ Mean	antigen	ctrl-	Δ Mean	antigen	ctrl-
CJB111	V	370	72 %	40%	63	57%	40%
COH1	III	305	76 %	10%	130	3%	10%
2603	V	82	22 %	34%	313	83 %	34%
7357b-	Ib	91	36%	34%	102	43%	34%
18RS21	II	0	15%	24%	268	84 %	24%
DK21	II	0	10%	21%	416	67 %	25%
A909	Ia	0	0%	14%			
O90	Ia	0	0%	0%			
H36B	Ib				105	34%	32%

Thus, inclusion of a non-AI protein in an immunogenic composition of the invention may provide increased protection a mammal.

The immunogenic compositions comprising *S. pneumoniae* AI polypeptides may further secondary SP protein antigens which include (a) any of the SP protein antigens disclosed in WO 02/077021 or U.S. provisional application _____, filed April 20, 2005 (Attorney Docket Number 002441.00154), (2) immunogenic portions of the antigens comprising at least 7 contiguous amino acids, (3) proteins comprising amino acid sequences which retain immunogenicity and which are at least 95% identical to these SP protein antigens (e.g., 95%, 96%, 97%, 98%, 99%, or 99.5% identical), and (4) fusion proteins, including hybrid SP protein antigens, comprising (1)-(3).

Alternatively, the invention may include an immunogenic composition comprising a first and a second Gram positive bacteria non-AI protein, wherein the polynucleotide sequence encoding the sequence of the first non-AI protein is less than 90% (i.e., less than 90, 88, 86, 84, 82, 81, 78, 76, 74, 72, 70, 65, 60, 55, 50, 45, 40, 35, or 30 percent) homologous than the corresponding sequence in the genome of the second non-AI protein.

The compositions of the invention may further comprise one or more additional non-Gram positive bacterial antigens, including additional bacterial, viral or parasitic antigens. The compositions of the invention may further comprise one or more additional non-GBS antigens, including additional bacterial, viral or parasitic antigens.

In another embodiment, the GBS antigen combinations of the invention are combined with one or more additional, non-GBS antigens suitable for use in a vaccine designed to protect elderly or immunocompromised individuals. For example, the GBS antigen combinations may be combined with an antigen derived from the group consisting of *Enterococcus faecalis*, *Staphylococcus aureus*, *Staphylococcus epidermis*, *Pseudomonas aeruginosa*, *Legionella pneumophila*, *Listeria monocytogenes*, *Neisseria meningitides*, influenza, and Parainfluenza virus ('PIV').

Where a saccharide or carbohydrate antigen is used, it is preferably conjugated to a carrier protein in order to enhance immunogenicity {e.g. Ramsay *et al.* (2001) *Lancet* 357(9251):195-196; Lindberg (1999) *Vaccine* 17 Suppl 2:S28-36; Buttery & Moxon (2000) *J R Coll Physicians Lond* 34:163-168; Ahmad & Chapnick (1999) *Infect Dis Clin North Am* 13:113-133, vii.; Goldblatt (1998) *J. Med. Microbiol.* 47:563-567; European patent 0 477 508; US Patent No. 5,306,492; International patent application WO98/42721; *Conjugate Vaccines* (eds. Cruse *et al.*) ISBN 3805549326, particularly vol. 10:48-114; and Hermanson (1996) *Bioconjugate Techniques* ISBN: 0123423368 or 012342335X}. Preferred carrier proteins are bacterial toxins or toxoids, such as diphtheria or tetanus toxoids. The CRM₁₉₇ diphtheria toxoid is particularly preferred {*Research Disclosure*, 453077 (Jan 2002)}. Other carrier polypeptides include the *N.meningitidis* outer membrane protein (EP-A-0372501), synthetic peptides (EP-A-0378881; EP-A-0427347), heat shock proteins (WO 93/17712; WO 94/03208), pertussis proteins (WO 98/58668; EP A 0471177), protein D from *H.influenzae* (WO 00/56360), cytokines (WO 91/01146), lymphokines, hormones, growth factors, toxin A or B from *C.difficile* (WO00/61761), iron-uptake proteins (WO01/72337), *etc.* Where a mixture comprises capsular saccharides from both serogroups A and C, it may be preferred that the ratio (w/w) of MenA saccharide:MenC saccharide is greater than 1 (e.g. 2:1, 3:1, 4:1, 5:1, 10:1 or higher). Different saccharides can be conjugated to the same or different type of carrier protein. Any suitable conjugation reaction can be used, with any suitable linker where necessary.

Toxic protein antigens may be detoxified where necessary e.g. detoxification of pertussis toxin by chemical and/or genetic means.

Where a diphtheria antigen is included in the composition it is preferred also to include tetanus antigen and pertussis antigens. Similarly, where a tetanus antigen is included it is preferred also to include diphtheria and pertussis antigens. Similarly, where a pertussis antigen is included it is preferred also to include diphtheria and tetanus antigens.

Antigens in the composition will typically be present at a concentration of at least 1 µg/ml each. In general, the concentration of any given antigen will be sufficient to elicit an immune response against that antigen.

As an alternative to using protein antigens in the composition of the invention, nucleic acid encoding the antigen may be used {e.g. refs. Robinson & Torres (1997) *Seminars in Immunology* 9:271-283; Donnelly *et al.* (1997) *Annu Rev Immunol* 15:617-648; Scott-Taylor & Dalgleish (2000) *Expert Opin Investig Drugs* 9:471-480; Apostolopoulos & Plebanski (2000) *Curr Opin Mol Ther* 2:441-447; Ilan (1999) *Curr Opin Mol Ther* 1:116-120; Dubensky *et al.* (2000) *Mol Med* 6:723-732; Robinson & Pertmer (2000) *Adv Virus Res* 55:1-74; Donnelly *et al.* (2000) *Am J Respir Crit Care Med* 162(4 Pt 2):S190-193; and Davis (1999) *Mt. Sinai J. Med.* 66:84-90}. Protein components of the compositions of the invention may thus be replaced by nucleic acid (preferably DNA e.g. in the form of a plasmid) that encodes the protein.

Definitions

PCT/US2005/027239

The term "comprising" means "including" as well as "consisting" *e.g.* a composition "comprising" X may consist exclusively of X or may include something additional *e.g.* X + Y.

The term "about" in relation to a numerical value *x* means, for example, $x \pm 10\%$.

References to a percentage sequence identity between two amino acid sequences means that, when aligned, that percentage of amino acids are the same in comparing the two sequences. This alignment and the percent homology or sequence identity can be determined using software programs known in the art, for example those described in section 7.7.18 of *Current Protocols in Molecular Biology* (F.M. Ausubel *et al.*, eds., 1987) Supplement 30. A preferred alignment is determined by the Smith-Waterman homology search algorithm using an affine gap search with a gap open penalty of 12 and a gap extension penalty of 2, BLOSUM matrix of 62. The Smith-Waterman homology search algorithm is disclosed in Smith & Waterman (1981) *Adv. Appl. Math.* 2: 482-489.

The invention is further illustrated, without limitation, by the following examples.

EXAMPLE 1: Binding of an Adhesin Island surface protein, GBS 80, to Fibrinogen and Fibronectin.

This example demonstrates that an Adhesin Island surface protein, GBS 80 can bind to fibrinogen and fibronectin.

An enzyme-linked immunosorbent assay (ELISA) was used to analyse the *in vitro* binding ability of recombinant GBS 80 to immobilized extra-cellular matrix (ECM) proteins but not to bovine serum albumin (BSA). Microtiter plates were coated with ECM proteins (fibrinogen, fibronectin, laminin, collagen type IV) and binding assessed by adding varying concentrations of a recombinant form of GBS 80, over-expressed and purified from *E. coli* (FIGURE 5A). Plates were then incubated sequentially with a) mouse anti-GBS 80 primary antibody; b) rabbit anti-mouse AP-conjugated secondary antibody; c) pNPP colorimetric substrate. Relative binding was measured by monitoring absorbance at 405 nm, using 595 nm as a reference wavelength. Figure 5b shows binding of recombinant GBS 80 to immobilized ECM proteins (1 μ g) as a function of concentration of GBS 80. BSA was used as a negative control. Data points represent the means of OD₄₀₅ values \pm standard deviation for 3 wells.

Binding of GBS 80 to the tested ECM proteins was found to be concentration dependent and exhibited saturation kinetics. As is also evident from FIGURE 5, binding of GBS 80 to fibronectin and fibrinogen was greater than binding to laminin and collagen type IV at all the concentrations tested.

EXAMPLE 2: GBS 80 is required for surface localization of GBS 104.

This example demonstrates that co-expression of GBS 80 is required for surface localization of GBS 104.

The polycistronic nature of the Adhesin Island I mRNA was investigated through reverse transcriptase-PCR (RT-PCR) analysis employing primers designed to detect transcripts arising from contiguous genes. Total RNA was isolated from GBS cultures grown to an optical density at 600 nm

(OD₆₀₀) of 0.3 in THB (Todd-Hewitt broth) by the RNeasy Total RNA isolation method (Qiagen) according to the manufacturer's instructions. The absence of contaminating chromosomal DNA was confirmed by failure of the gene amplification reactions to generate a product detectable by agarose gel electrophoresis, in the absence of reverse transcriptase. RT-PCR analysis was performed with the Access RT-PCR system (Promega) according to the manufacturer's instructions, employing PCR cycling temperatures of 60°C for annealing and 70°C for extension. Amplification products were visualized alongside 100-bp DNA markers in 2% agarose gels after ethidium bromide staining.

FIGURE 5 shows that all the genes are co-transcribed as an operon. A schematic of the AI-1 operon is shown above the agarose gel analysis of the RT-PCR products. Large rectangular arrows indicate the predicted transcript direction. Primer pairs were selected such as "1-4" cross the 3' finish-5' start of successive genes and overlap each gene by at least 200 bp. Additionally, "1" crosses a putative rho-independent transcriptional terminator. "5" is an internal GBS 80 control and "6" is an unrelated control from a highly expressed gene. Lanes: "a": RNA plus RTase enzyme; "b" RNA without RTase; "c": genomic DNA control.

In the effort to elucidate the functions of the AI-1 proteins, in frame deletions of all of the genes within the operon have been constructed and the resulting mutants characterized with respect to surface exposure of the encoded antigens (see FIGURE 8).

Each in-frame deletion mutation was constructed by splice overlap extension PCR (SOE-PCR) essentially as described by Horton et al. [Horton R. M., Z. L. Cai, S. N. Ho, L. R. Pease (1990) *Biotechniques* 8:528-35] using suitable primers and cloned into the temperature sensitive shuttle vector pJRS233 to replace the wild type copy by allelic exchange [Perez-Casal, J., J. A. Price, et al. (1993) *Mol Microbiol* 8(5): 809-19.]. All plasmid constructions utilized standard molecular biology techniques, and the identities of DNA fragments generated by PCR were verified by sequencing. Following SOE-PCR, the resulting mutant DNA fragments were digested with XhoI and EcoRI, and ligated into a similarly digested pJRS233. The resulting vectors were introduced by electroporation into the chromosome of 2603 and COH1 GBS strains in a three-step process, essentially as described in Framson et al. [Framson, P. E., A. Nittayajarn, J. Merry, P. Youngman, and C. E. Rubens. (1997) *Appl. Environ. Microbiol.* 63(9):3539-47]. Briefly, the vector pJRS233 contains an *erm* gene encoding erythromycin resistance and a temperature-sensitive gram-positive replicon that is active at 30°C but not at 37°C. Initially, the constructs are electroporated into GBS electro-competent cells prepared as described by Framson et al., and transformants containing free plasmid are selected by their ability to grow at 30°C on Todd-Hewitt Broth (THB) agar plates containing 1 µg/ml erythromycin. The second step includes a selection step for strains in which the plasmid has integrated into the chromosome via a single recombination event over the homologous plasmid insert and chromosome sequence by their ability to grow at 37°C on THB agar medium containing 1 mg/ml erythromycin. In the third step, GBS cells containing the plasmid integrated within the chromosome (integrants) are serially passed in broth culture in the absence of antibiotics at 30°C. Plasmid excision

from the chromosome via a second recombination event over the duplicated target gene sequence either completed the allelic exchange or reconstituted the wild-type genotype. Subsequent loss of the plasmid in the absence of antibiotic selection pressure resulted in an erythromycin-sensitive phenotype. In order to assess gene replacement a screening of erythromycin-sensitive colonies was performed by analysis of the target gene PCR amplicons.

FIGURE 7 reports a schematic of the IS-1 operon for each knock-out strain generated, along with the deletion position within the amino acidic sequence. Most data presented here concern the COH1 deletion strains, in which the expression of each of the antigens is higher by DNA microarray analysis (data not shown) as well as detectable by FACS analysis (see FIGURE 8). The double mutant in 2603 Δ 80, Δ 104 double mutant was constructed by sequential allelic exchanges of the shown alleles.

Immunization protocol

Immune sera for FACS experiments were obtained as follows.

Groups of 4 CD-1 outbred female mice 6-7 weeks old (Charles River Laboratories, Calco Italy) were immunized with the selected GBS antigens, (20 μ g of each recombinant GBS antigen), suspended in 100 μ l of PBS. Each group received 3 doses at days 0, 21 and 35. Immunization was performed through intra-peritoneal injection of the protein with an equal volume of Complete Freund's Adjuvant (CFA) for the first dose and Incomplete Freund's Adjuvant (IFA) for the following two doses. In each immunization scheme negative and positive control groups are used. Immune response was monitored by using serum samples taken on day 0 and 49.

FACS analysis

Preparation of paraformaldehyde treated GBS cells and their FACS analysis were carried out as follows.

GBS serotype COH1 strain cells were grown in Todd Hewitt Broth (THB; Difco Laboratories, Detroit, Mich.) to OD_{600nm} = 0.5. The culture was centrifuged for 20 minutes at 5000 rpm and bacteria were washed once with PBS, resuspended in PBS containing 0.05% paraformaldehyde, and incubated for 1 hours at 37 °C and then overnight at 4°C. 50 μ l of fixed bacteria (OD₆₀₀ 0.1) were washed once with PBS, resuspended in 20 μ l of Newborn Calf Serum, (Sigma) and incubated for 20 min. at room temperature. The cells were then incubated for 1 hour at 4°C in 100 μ l of preimmune or immune sera, diluted 1:200 in dilution buffer (PBS, 20% Newborn Calf Serum, 0.1% BSA). After centrifugation and washing with 200 μ l of washing buffer (0.1% BSA in PBS), samples were incubated for 1 hour at 4°C with 50 μ l of R-Phicoerytrin conjugated F(ab)₂ goat anti-mouse IgG (Jackson ImmunoResearch Laboratories; Inc.), diluted 1:100 in dilution buffer. Cells were washed with 200 μ l of washing buffer and resuspended in 200 μ l of PBS. Samples were analysed using a FACS Calibur apparatus (Becton Dickinson, Mountain View, Calif.) and data were analyzed using the Cell Quest Software (Becton Dickinson). A shift in mean fluorescence intensity of > 75 channels compared to preimmune sera from the same mice was considered positive. This cutoff

was determined from the mean plus two standard deviations of shifts obtained with control sera raised against mock purified recombinant proteins from cultures of *E. coli* carrying the empty expression vector and included in every experiment. Artifacts due to bacterial lysis were excluded using antisera raised against 6 different known cytoplasmic proteins all of which were negative

5 FACS data on COH1 single KO mutants for GBS 104 and GBS 80 indicated that GBS 80 is required for surface localization of GBS 104.

As shown in FIGURE 8, GBS 104 is not surface exposed in the $\Delta 80$ strain (second column, bottom), but is present in the whole protein extracts (see FIGURE 10). Mean shift values suggest that GBS 104 is partially responsible for GBS 80 surface exposure (Mean shift of GBS 80 is reduced to
10 ~60% wild-type levels in $\Delta 104$), and that GBS 80 is over-expressed in the complemented strain (mean shift value ~200% wild-type level). The $\Delta 80/p$ GBS 80 strain contains the GBS 80 orf cloned in the shuttle-vector pAM401 (Wirth, R., F. Y. An, et al. (1986). J Bacteriol 165(3): 831-6). The vector alone does not alter the secretion pattern of GBS 104 (right column). FACS was performed on mid-log fixed bacteria with mouse polyclonal antibodies as indicated at left. Black peak is pre-immune
15 sera, colored peaks are sera from immunized animals.

EXAMPLE 3: Deletion of GBS 80 causes attenuation *in vivo*.

This example demonstrates that deletion of GBS 80 causes attenuation *in vivo*, suggesting that this protein contributes to bacterial virulence.

20 By using a mouse animal model, we studied the role of GBS 80 and GBS 104 in the virulence of *S. agalactiae*.

Groups of ten outbred female mice 5-6 week weeks old (Charles River Laboratories, Calco Italy) were inoculated intraperitoneally with different dilutions of the mutant strains and LD50 (lethal dose 50) were calculated according to the method of Reed and Muench [Reed, L. J. and H. Muench (1938).The American Journal of Hygiene 27(3): 493-7]. As presented in the table below, the number
25 of colony forming units (cfu) counted for both the $\Delta 80$ and the $\Delta 80, \Delta 104$ double mutants is about 10 fold higher when compared to the wild type strain suggesting that inactivation of GBS 80 but not GBS 104 is responsible for an attenuation in virulence. This finding indicates that GBS 80 gene in the AI-1 might contribute to virulence.

30 Table Lethal dose 50% analysis of AI-1 mutants in the 2603 strain background. LD50s were performed by IP injection of female CD1 mice at an age of 5-6 weeks. LD50s were calculated by the method of Reed and Muench (8).

GBS strain	LD ₅₀ , cfu	Number of Experiments
Wild Type 2603	2×10^8	4
$\Delta 104$ mutant	$\sim 2 \times 10^8$	1
$\Delta 80$ mutant	2.6×10^9	3
$\Delta 80, \Delta 104$ double mutant	$\sim 2 \times 10^9$	1

EXAMPLE 4: Effect of Adhesin Island Sortase Deletions on Surface Antigen Presentation

This example demonstrates the effect of adhesin island sortase deletions on surface antigen presentation.

FACS analysis results set forth in FIGURE 9 show that a deletion in sortase SAG0648 prevented GBS 104 from reaching the surface and slightly reduced the surface exposure of GBS 80 (fourth panel; mean shift value ~60% wild-type COH1). In the double sortase knock-out strain, neither antigen was surface exposed (far right panel). Either sortase alone was sufficient for GBS 80 to arrive at the bacterial surface (third and fourth columns, top). No effect was seen on surface exposure of antigens GBS 80 or GBS 104 in the Δ GBS 52 strain. Antibodies derived from purified GBS 52 were either non-specific or were FACS negative for GBS 52 (data not shown). FACS analysis was performed as described above (see EXAMPLE 2).

As shown in FIGURE 10, inactivation of GBS 80 has no effect on GBS 104 expression as much as GBS 104 knock out doesn't change the total amount GBS 80 expressed. The Western blot of whole protein extracts (strains noted above lanes) probed with anti-GBS 80 antisera is shown in panel A. Arrow indicates expected size of GBS 80 (60 kDa). GBS 80 antibodies recognize a doublet, the lower band is not present in Δ GBS 80 strains. Panel B shows a Western blot of whole protein extracts probed with anti-GBS 104 antisera. Arrow indicates expected size of GBS 104 (99.4 kDa). Protein extracts were prepared from the same bacterial cultures used for FACS (FIGURES 8 and 9). In conclusion, although GBS 104 does not arrive at the surface in the Δ 80 strain by FACS (FIGURE 8, second column), it is present at approximately wild-type levels in the whole protein preps (B, second lane). Approximately 20 μ g of each protein extract was loaded per lane.

Western-blot analysis

Aliquots of total protein extract mixed with SDS loading buffer (1x: 60 mM TRIS-HCl pH 6.8, 5% w/v SDS, 10% v/v glycerol, 0.1% Bromophenol Blue, 100 mM DTT) and boiled 5 minutes at 95° C, were loaded on a 12.5% SDS-PAGE precast gel (Biorad). The gel is run using a SDS-PAGE running buffer containing 250 mM TRIS, 2.5 mM Glycine and 0.1 %SDS. The gel is electroblotted onto nitrocellulose membrane at 200 mA for 60 minutes. The membrane is blocked for 60 minutes with PBS/0.05 % Tween-20 (Sigma), 10% skimmed milk powder and incubated O/N at 4° C with PBS/0.05 % Tween 20, 1% skimmed milk powder, with the appropriate dilution of the sera. After washing twice with PBS/0.05 % Tween, the membrane is incubated for 2 hours with peroxidase-conjugated secondary anti-mouse antibody (Amersham) diluted 1:4000. The nitrocellulose is washed three times for 10 minutes with PBS/0.05 % Tween and once with PBS and thereafter developed by Opti-4CN Substrate Kit (Biorad).

Example 5: Binding of Adhesin Island proteins to epithelial cells and effect of Adhesin Island proteins on capacity of GBS to adhere to epithelial cells.

This example illustrates the binding of AI proteins to epithelial cells and the effect of AI proteins on the capacity of GBS to adhere to epithelial cells.

Applicants analysed whether recombinant AI surface proteins GBS 80 or GBS 104 would demonstrate binding to various epithelial cells in a FACS analysis. Applicants also analysed whether

deletion of AI surface proteins GBS 80 or GBS 104 would effect the capacity of GBS to adhere to and invade ME180 cervical epithelial cells.

As shown in Figure 28, deletion of GBS 80 sequence from GBS strain isolate 2603 (serotype V) did not affect the capacity of the mutated GBS to adhere to and invade ME180 cervical epithelial cells. Here ME180 cervical carcinoma epithelial cells were infected with wild type GBS 2603 or GBS 2603 Δ 80 isogenic mutant. After two hours of infection, non-adherent bacteria were washed off and infection prolonged for a further two hours and four hours. In invasion experiments, after each time point, was followed by a two hour antibiotic treatment. Cells were then lysed with 1% saponin and lysates plated on TSA plates. As shown in Figure 28, there was little difference between the percent invasion or percent adhesion of wild type and mutant strains up to the four hour time point.

Figure 30 repeats this experiment with both Δ 104 and Δ 80 mutants from a different strain isolate. Here, ME180 cervical carcinoma epithelial cells were infected with GBS strain isolate COH (serotype III) wild type or COH1 Δ GBS 104 or COH1 Δ 80 isogenic mutant. After one hour of infection, non-adherent bacteria were washed off and the cells were lysed with 1% saponin. The lysates were plated on TSA plates. As shown in Figure 30, while there was little difference in the percent invasion, there was a significant decrease in the percent association of the Δ 104 mutant compared to both the wild type and Δ 80 mutant.

The affect of AI surface proteins on the ability of GBS to translocate through an epithelial monolayer was also analysed. As shown in Figure 31, a GBS 80 knockout mutant strain partially loses the ability to translocate through an epithelial monolayer. Here epithelial monolayers were inoculated with wildtype or knockout mutant in the apical chamber of a transwell system for two hours and then non-adherent bacteria were washed off. Infection was prolonged for a further two and four hours. Samples were taken from the media of the basolateral side and the number of colony forming units measured. Transepithelial electrical resistance measured prior to and after infection gave comparable values, indicating the maintenance of the integrity of the monolayer. By the six hour time point, the Δ 80 mutants demonstrated a reduced percent transcytosis.

A similar experiment was conducted with GBS 104 knock out mutants. Here, as shown in Figure 22, the Δ 104 mutants also demonstrated a reduced percent transcytosis, indicating that the mutant strains translocate through an epithelial monolayer less efficiently than their isogenic wild type counterparts.

Applicants also studied the effect of AI proteins on the capacity of a GBS strain to invade J774 macrophage-like cells. Here, J774 cells were infected with GBS COH1 wild type or COH1 Δ GBS104 or COH1 Δ GBS80 isogenic mutants. After one hour of infection, non-adherent bacteria were washed off and intracellular bacteria were recovered at two, four and six hours post antibiotic treatment. At each time point, cells were lysed with 0.25% Triton X-100 and lysates plated on TSA plates. As shown in Figure 32, the Δ 104 mutant demonstrated a significantly reduced percent invasion compared to both the wild type and Δ 80 mutant.

Example 6: Hyperoligomeric structures comprising AI surface proteins GBS 80 and GBS 104.

This example illustrates hyperoligomeric structures comprising AI surface proteins GBS 80 and GBS 104. A GBS isolate COH1 (serotype III) was adapted to increase expression of GBS 80.

Figure 34 presents a regular negative stain electron micrograph of this mutant; no pilus or

hyperoligomeric structures are distinguishable on the surface of the bacteria. When the EM stain is based on anti-GBS 80 antibodies labelled with 10 or 20 nm gold particles, the presence of GBS 80 throughout the hyperoligomeric structure is clearly indicated (Figures 36, 37 and 38). EM staining against GBS 104 (anti-GBS 104 antibodies labelled with 10 nm gold particles) also reveals the presence of GBS 104 primarily on or near the surface of the bacteria or potentially associated with bacterial peptidoglycans (Figure 39). Analysis of this same strain (over-expressing GBS 80) with a combination of both anti-GBS 80 (using 20 nm gold particles) and anti-GBS 104 (using 10 nm gold particles) reveals the presence of GBS 104 on the surface and within the hyperoligomeric structures (see Figures 40 and 41).

Example 7: GBS 80 is necessary for polymer formation and GBS 104 and sortase SAG0648 are necessary for efficient pili assembly

This example demonstrates that GBS 80 is necessary for formation of polymers and that GBS 104 and sortase SAG0648 are necessary for efficient pili assembly. GBS 80 and GBS 104 polymeric assembly was systematically analyzed in Coh1 strain single knock out mutants of each of the relevant coding genes in AI-1 (GBS 80, GBS 104, GBS 52, sag0647, and sag0648). Figure 41 provides Western blots of total protein extracts (strains noted above lanes) probed with either anti-GBS 80 (left panel) sera or anti-GBS 104 sera (right panel) for each of these Coh1 and Coh1 knock out strains. (Coh1, wild type Coh1; Δ 80, Coh1 with GBS 80 knocked out; Δ 104, Coh1 with GBS 104 knocked out; Δ 52, Coh1 with GBS 52 knocked out; Δ 647, Coh1 with SAG0647 knocked out; Δ 648, Coh1 with SAG0648 knocked out; Δ 647-8, Coh1 with SAG0647 and SAG0648 knocked out; Δ 80/pGBS80, Coh1 with GBS 80 knocked out but complemented with a high copy number plasmid expressing GBS 80. Asterisks identify the monomer of GBS 80 and GBS 104.)

The smear of immunoreactive material observed in the wild type strain, along with its disappearance in Δ 80 and Δ 104 mutants, is consistent with the notion that such high molecular weight structures are composed of covalently linked (SDS-resistant) GBS 80 and GBS 104 subunits. The immunoblotting with both anti-GBS 80 (α -GBS 80) and anti-GBS 104 (α -GBS 104) revealed that deletion of sortase SAG0648 also interferes with the assembly of high molecular weight species, whereas the knock out mutant of the second sortase (SAG0647), even if somehow reduced, still maintains the ability to form polymeric structures.

Total extracts from GBS were prepared as follows. Bacteria were grown in 50 ml of Todd-Hewitt broth (Difco) to an OD_{600nm} of 0.5-0.6 and successively pelleted. After two washes in PBS the pellet was resuspended and incubated 3 hours at 37°C with mutanolisin. Cells were then lysed with at

least three freezing-thawing cycles in dry ice and a 37°C bath. The lysate was then centrifuged to eliminate the cellular debris and the supernatant was quantified. Approximately 40 µg of each protein extract was separated on SDS-PAGE. The gel was then subjected to immunoblotting with mice antisera and detected with chemiluminescence.

Example 8: GBS 80 is polymerized by an AI-2 sortase

This example illustrates that GBS 80 can be polymerized not only by AI-1 sortases, but also by AI-2 sortases. Figure 42 shows total cell extract immunoblots of GBS 515 strain, which lacks AI-1. The left panel, where an anti-GBS 67 sera was used, shows that GBS 67 from AI-2 is assembled into high-molecular weight-complexes, suggesting the formation of a second type of pilus. The same high molecular structure is observed when GBS 80 is highly expressed by reintroducing the gene within a plasmid (pGBS 80). By using anti-GBS 80 (right panel) sera on the same extracts, again it is observed that, with GBS 80 over expression (515/pGBS 80), a high-molecular weight structure is assembled. This implies that, in the absence of AI-1 sortases, AI-2 sortases (SAG1405 and SAG1406) can complement the lacking function, still being able to assemble GBS 80 in a pilus structure.

Example 9: Coh1 produces a high molecular weight molecule, the GBS 80 pilin

This example illustrates that Coh1 produces a high molecular weight molecule, greater than 1000 kDa, which is the GBS 80 pilin. Figure 43 provides silver-stained electrophoretic gels that show that Coh1 produces two macromolecules. One of these macromolecules disappears in the Coh1 GBS 80 knock out cells, but does not disappear in the Coh1 GBS 52 knock out mutant cells. The last two lanes on the right were loaded with 15 times the amount loaded in the other lanes. This was done in order to be able to count the bands. By doing this, a conservative size estimate of the top bands was calculated by starting at 240 kDa and considering each of 14 higher bands as the result of consecutive additions of a GBS 80 monomer.

Coh1, wild type Coh1; Δ80, Coh1 cells with GBS 80 knocked out; Δ52, Coh1 cells with GBS 52 knocked out; Δ80/pGBS 80, Coh1 cells with GBS 80 knocked out and complemented with a high copy number construct expressing GBS 80.

Example 10. GBS 52 is a minor component of the GBS pilus

This example illustrates that GBS 52 is present in the GBS pilus and is a minor component of the pilus. Figure 45 shows an immunoblot of total cell extracts from a GBS Coh1 strain and a GBS Coh1 strain knocked out for GBS 52 (Δ52). The total cell extracts were immunoblotted anti-GBS 80 antisera (left) and anti-GBS 52 antisera (right). Immunoblotting was performed using a 3-8% Tris-acetate polyacrylamide gel (Invitrogen) which provided excellent separation of large molecular weight proteins (see figure 41). When the gel was incubated with anti-GBS 80 sera, the bands from the Coh1 wild-type strain appeared shifted when compared to the Δ52 mutant. This observation

indicated a different size of the pilus polymeric components in the two strains. When the same gel was stripped and incubated with anti-GBS 52 sera the high-molecular subunits in the Coh1 wild-type strain showed similar molecular size of those in the correspondent lane in the left panel. These findings confirmed that GBS 52 is indeed associated with GBS 80 macro-molecular structures but represents a minor component of the GBS pilus.

Example 11: Pilus structures are present in the supernatant of GBS bacterial cultures

This example illustrates that the pilus structure assembled in Coh1 GBS is present in the supernatant of a bacterial cell culture. Figure 46 shows an immunoblot where the protein extract of the supernatant from cultures of different GBS mutant strains (117 = Coh1 GBS 80 knockout; 159 = Coh1 GBS 104 knockout; 202 = Coh1 GBS 52 knockout; 206 = Coh1 GBS sag0647 knockout; 208 = Coh1 GBS sag0648 knockout; 197 = Coh1 GBS sag0647/sag0648 knockout; 179 = Coh1 GBS 80 knockout complemented with a high copy plasmid expressing GBS 80). GBS 80 antisera detects the presence of pilus structures in the appropriate Coh1 strains.

The protein extract was prepared as follows. Bacteria were grown in THB to an OD_{600nm} of 0.5-0.6 and the supernatant was separated from the cells by centrifugation. The supernatant was then filtered (\varnothing 0.2 μ m) and 1 ml was added with 60% TCA for protein precipitation. GBS pili were also extracted from the fraction of surface-exposed proteins in Coh1 strain and its GBS 80 knock out mutant as described hereafter. Bacteria were grown to an OD_{600nm} of 0.6 in 50 ml of THB at 37°C. Cells were washed once with PBS and the pellet was then resuspended in 0.1 M KPO4 pH 6.2, 40% sucrose, 10 mM MgCl₂, 400U/ml mutanolysin and incubated 3 hours at 37°C. Protoplasts were separated by centrifugation and the supernatant was recovered and its protein content measured.

In order to study the dynamics of pilus production during different growth phases, 1 ml supernatant of a culture at different OD_{600nm} was TCA precipitated and loaded onto a 3-8% SDS-PAGE as described before. Figure 47 shows the corresponding Western blot with GBS 80 anti-sera. The first group of lanes (left five sample lanes) refer to a Coh1 strain growth (OD_{600nm} are noted above the lanes) whereas the second group of lanes (right five samples) are from a GBS 80 knock out strain over expressing GBS 80. The experiment shows that pilus macromolecular structures can be found in the supernatant in all of the growth phases tested.

Example 12: In GBS strain Coh1, only GBS 80 and a sortase (sag0647 or sag0648) is required for polymerization

This example describes requirements for pilus formation in Coh1. Figure 48 shows a Western blot of total protein extracts (prepared as described before) using anti-GBS 80 sera on Coh1 clones. (Coh1, wild type Coh1; Δ 104, Coh1 knocked out for GBS 104, Δ 647, Coh1 knocked out for sag0647, Δ 648, Coh1 knocked for sag0648, Δ 647-8, Coh1 knocked out for sag0647 and sag0648; 515, wild

type bacterial strain 515, which lacks an AI-1; p80 a high copy number plasmid which expresses GBS 80.) The data show that only the double sortase mutant is unable to polymerize GBS 80 indicating that the 'conditio sine qua non' for pilus polymerization is the co-existence of GBS 80 with at least one sortase. This result leads to a reasonable assumption that SAG1405 and SAG1406 are responsible for polymerization in this strain.

Example 13: GBS 80 can be expressed in *L. lactis* under its own promoter and terminator sequences

This example demonstrates that *L. lactis*, a non-pathogenic bacterium, can express GBS AI polypeptides such as GBS 80. *L. lactis* M1363 (*J. Bacteriol.* 154 (1983):1-9) was transformed with a construct encoding GBS 80. Briefly, the construct was prepared by cloning a DNA fragment containing the gene coding for GBS 80 under its own promoter and terminator sequences into plasmid pAM401 (a shuttle vector for *E. coli* and other Gram positive bacteria; *J. Bacteriol.* 163 (1986):831-836). Total extracts of the transformed bacteria in log phase were separated on SDS-PAGE, transferred to membranes, and incubated with antiserum against GBS 80. A polypeptide corresponding to the molecular weight of GBS 80 was detected in the lanes containing total extracts of *L. lactis* transformed with the GBS 80 construct. See Figures 133A and 133B, lanes 6 and 7. This same polypeptide was not detected in the lane containing total extracts of *L. lactis* not transformed with the GBS 80 construct, lane 9. This example shows that *L. lactis* can express GBS 80 under its own promoter and terminator.

Example 14: *L. lactis* modified to express GBS AI-1 under the GBS 80 promoter and terminator sequences expresses GBS 80 in polymeric structures

This example demonstrates the ability of *L. lactis* to express GBS AI-1 polypeptides and to incorporate at least some of the polypeptides into oligomers. *L. lactis* was transformed with a construct containing the genes encoding GBS AI-1 polypeptides. Briefly, the construct was prepared by cloning a DNA fragment containing the genes for GBS 80, GBS 52, SAG0647, SAG0648, and GBS 104 under the GBS 80 promoter and terminator sequences into construct pAM401. The construct was transformed into *L. lactis* M1363. Total extracts of log phase transformed bacteria were separated on reducing SDS-PAGE, transferred to membranes, and incubated with antiserum against GBS 80. A polypeptide with a molecular weight corresponding to the molecular weight of GBS 80 was detected in the lanes containing *L. lactis* transformed with the GBS AI-1 encoding construct. See Figure 134, lane 2. In addition, the same lane also showed immunoreactivity of polypeptides having higher molecular weights than the polypeptide having the molecular weight of GBS 80. These higher molecular weight polypeptides are likely oligomers of GBS 80. Oligomers of similar molecular

weights were also observed on a Western blot of the culture supernatant of the transformed *L. lactis*. See lane 4 of Figure 135. Thus, this example shows that *L. lactis* transformed to express GBS AI-1 can efficiently polymerize GBS 80 in the form of a pilus. This pilus structure can likely be purified from either the cell culture supernatant or cell extracts.

5

Example 15: Cloning and Expression of *S. pneumoniae* Sp0462

This example describes the production of a clone encoding a Sp0462 polypeptide and expression of the clone. To produce a clone encoding Sp0462, the open reading frame encoding Sp0462 was amplified using primers that annealed within the full-length Sp0462 open reading frame sequence. Figure 150A provides a 893 amino acid sequence of Sp0462. The primers used to produce a clone encoding the Sp0462 polypeptide are shown in Figure 150B. These primers annealed to the nucleotide sequences encoding the amino acid residues indicated by underlining in Figure 150A. Amplification of the open reading frame encoding Sp0462 using these primers produced the amplicon shown at lane 2 of the agarose gel provided in Figure 160. The Sp0462 clone encodes amino acid residues 38-862 of the 893 amino acid residue Sp0462 protein; the italicized residues in Figure 150A were eliminated. Figure 151A provides a schematic depiction of the recombinant Sp0462 polypeptide. Figure 151B shows a schematic depiction of the full-length Sp0462 polypeptide. Both the recombinant Sp0462 encoded by the clone and the full-length Sp0462 protein have two collagen binding protein type B (Cna B) domains and a von Hillebrand factor A (vWA) domain. The cloned recombinant Sp0462 lacks the LPXTG motif present in the full-length Sp0462 protein. Western blot analysis for expression of the Sp0462 clone did not result in detection of polypeptides with serum obtained from *S. pneumoniae*-infected patients (Figure 152A) or GBS 80 antiserum (Figure 152B).

Example 16: Cloning and Expression of *S. pneumoniae* Sp0463

This example describes the production of a clone encoding a Sp0463 polypeptide and detection of recombinant Sp0463 polypeptide expressed from the clone. To produce a clone encoding Sp0463, the open reading frame encoding Sp0463 was amplified using primers that annealed within the full-length Sp0463 open reading frame sequence. Figure 153A provides a 665 amino acid sequence of Sp0463. The primers used to produce the clone encoding Sp0463 polypeptide are shown in Figure 153B. These primers annealed to the nucleotide sequences encoding the amino acid residues indicated by underlining in Figure 153A. Amplification of the open reading frame encoding Sp0463 using these primers produced the amplicon shown at lane 3 of the agarose gel provided in Figure 160. The Sp0463 clone encodes amino acid residues 23-627 of the 665 amino acid residue Sp0463 protein; the italicized residues in Figure 153A were eliminated. Figure 154A provides a schematic depiction of the recombinant Sp0463 polypeptide. Figure 154B shows a schematic depiction of the full-length Sp0463 polypeptide. Both the recombinant Sp0463 encoded by the clone and the full-length Sp0463 protein have a Cna B domain and an E box motif. The cloned recombinant

Sp0463 lacks the LPXTG motif present in the full-length Sp0463 protein. Expression of the Sp0463 clone resulted in the detection of a 60 kD polypeptide, the expected molecular weight of the recombinant Sp0463 polypeptide, by Western blot analysis. See Figure 155.

5 Example 17: Cloning and Expression of *S. pneumoniae* Sp0464

This example describes the production of a clone encoding a Sp0464 polypeptide and detection of recombinant Sp0464 polypeptide expressed from the clone. To produce a clone encoding Sp0464, the open reading frame encoding Sp0464 was amplified using primers that annealed either within the full-length Sp0464 open reading frame sequence. Figure 157A provides a 393 amino acid
10 sequence of Sp0464. The primers used to produce a clone encoding the Sp0464 polypeptide are shown in Figure 157B. These primers annealed to the nucleotide sequences encoding the amino acid residues indicated by underlining in Figure 157A. Amplification of the open reading frame encoding Sp0464 using these primers produced the amplicon shown at lane 4 of the agarose gel provided in Figure 160. The Sp0464 clone encodes amino acid residues 19-356 of the 393 amino acid residue
15 Sp0464 protein; the italicized residues in Figure 157A were eliminated. Figure 158A provides a schematic depiction of the recombinant Sp0464 polypeptide. Figure 158B shows a schematic depiction of the full-length Sp0464 polypeptide. Both the recombinant Sp0464 encoded by the clone and the full-length Sp0464 protein have two Cna B domains. The cloned recombinant Sp0464 lacks the LPXTG motif present in the full-length Sp0464 protein. Expression of the Sp0464 clone resulted
20 in the detection of a 38 kD polypeptide, the expected molecular weight of the recombinant Sp0464 polypeptide, by Western blot analysis. See Figure 159.

Example 18: Intranasal Immunization of Mice with Recombinant *L. lactis* Expressing GBS 80 and Subsequent Challenge

25 This example describes a method of intranasally immunizing mice using *L. lactis* that express GBS 80. Intranasal immunization consisted of 3 doses at days 0, 14 and 28, each dose administered in three consecutive days. Each day, groups of 3 CD-1 outbred female mice 6-7 weeks old (Charles River Laboratories, Calco Italy) were immunized intranasally with 10^9 or 10^{10} CFU of the recombinant *Lactococcus lactis* suspended in 20 μ l of PBS. In each immunization scheme negative
30 (wild-type *L. lactis*) and positive (recombinant GBS80) control groups were used. The immune response of the dams was monitored by using serum samples taken on day 0 and 49. The female mice were bred 2-7 days after the last immunization (at approximately $t = 36 - 37$), and typically had a gestation period of 21 days. Within 48 hours of birth, the pups were challenged via I.P. with GBS in a dose approximately equal to an amount which would be sufficient to kill 90 % of immunized pups (as
35 determined by empirical data gathered from PBS control groups). The GBS challenge dose is preferably administered in 50ml of THB medium. Preferably, the pup challenge takes place at 56 to 61 days after the first immunization. The challenge inocula were prepared starting from frozen

cultures diluted to the appropriate concentration with THB prior to use. Survival of pups was monitored for 5 days after challenge.

Example 19: Subcutaneous Immunization of Mice with Recombinant *L. lactis* Expressing GBS 80 and Subsequent Challenge

This example describes a method of subcutaneous immunization mice using *L. lactis* that express GBS 80. Subcutaneous immunization consists of 3 doses at days 0, 14 and 28. Groups of 3 CD-1 outbred female mice 6-7 weeks old (Charles River Laboratories, Calco Italy) were injected subcutaneously with 10^9 or 10^{10} CFU of the recombinant *Lactococcus lactis* suspended in 100 μ l of PBS. In each immunization scheme, negative (wild-type *L. lactis*) and positive (recombinant GBS80) control groups were used. The immune response of the dams was monitored by using serum samples taken on day 0 and 49. The female mice were bred 2-7 days after the last immunization (at approximately $t = 36 - 37$), and typically had a gestation period of 21 days. Within 48 hours of birth, the pups were challenged via I.P. with GBS in a dose approximately equal to an amount which would be sufficient to kill 90 % of immunized pups (as determined by empirical data gathered from PBS control groups). The GBS challenge dose is preferably administered in 50ml of THB medium. Preferably, the pup challenge takes place at 56 to 61 days after the first immunization. The challenge inocula were prepared starting from frozen cultures diluted to the appropriate concentration with THB prior to use. Survival of pups was monitored for 5 days after challenge.

Example 20: Immunization of Mice with GAS AI polypeptides and Subsequent Intranasal Challenge

This example describes a method of immunizing mice with GAS AI polypeptides and subsequently intranasally challenging the mice with GAS bacteria. Groups of 10 CD1 female mice aged between 6 and 7 weeks are immunized with a combination of GAS antigens of the invention GAS 15, GAS 16, and GAS 18, (15 μ g of each recombinant antigen, derived from M1 strain SF370) or *L. lactis* expressing the M1 strain SF370 adhesin island, suspended in 100 μ l of suitable solution. Each group receives 3 doses at days 0, 21 and 45. Immunization is performed through subcutaneous or intraperitoneal injection for the GAS 15, GAS 16, GAS 18 protein combination. The protein combination is administered with an equal volume of Complete Freund's Adjuvant (CFA) for the first dose and Incomplete Freund's Adjuvant (IFA) for the following two doses. Immunization is performed intranasally for the *L. lactis* expressing the M1 strain SF370 adhesin island. In each immunization scheme negative and positive control groups are used.

The negative control group for the mice immunized with the GAS 15, GAS 16, GAS 18 protein combination included mice immunized with PBS. The negative control group for the mice immunized with *L. lactis* expressing the M1 strain SF370 adhesin island, included mice immunized

with either wildtype *L. lactis* or *L. lactis* transformed with the pAM401 expression vector lacking any cloned adhesin island sequence.

The positive control groups included mice immunized with purified M1 strain SF370 M protein.

- 5 Immunized mice are then anaesthetized with Zoletil and challenged intranasally with a 25 µL suspension containing 1.2×10^6 or 1.2×10^8 CFU of ISS 3348 in THB. Animals are observed daily and checked for survival.

Example 21: Active Maternal Immunization Assay

- 10 As used herein, an Active Maternal Immunization assay refers to an *in vivo* protection assay where female mice are immunized with the test antigen composition. The female mice are then bred and their pups are challenged with a lethal dose of GBS. Serum titers of the female mice during the immunization schedule are measured as well as the survival time of the pups after challenge.

15 Mouse immunization

- Specifically, groups of 4 CD-1 outbred female mice 6-8 weeks old (Charles River Laboratories, Calco Italy) are immunized with one or more GBS antigens, (20 µg of each recombinant GBS antigen), suspended in 100 µl of PBS. Each group receives 3 doses at days 0, 21 and 35. Immunization is performed through intra-peritoneal injection of the protein with an equal volume of Complete Freund's Adjuvant (CFA) for the first dose and Incomplete Freund's Adjuvant (IFA) for the following two doses. In each immunization scheme negative and positive control groups are used.

Immune response is monitored by using serum samples taken on day 0 and 49. The sera are analyzed as pools from each group of mice.

25 Active maternal immunization

- A maternal immunization/neonatal pup challenge model of GBS infection was used to verify the protective efficacy of the antigens in mice. The mouse protection study was adapted from Rodewald et al. (Rodewald et al. J. Infect. Diseases 166, 635 (1992)). In brief, CD-1 female mice (6-8 weeks old) were immunized before breeding, as described above. The mice received 20 µg of protein per dose when immunized with a single antigen and 60 µg of protein per dose (15 µg of each antigen) when immunized with the combination of antigens. Mice were bred 2-7 days after the last immunization. Within 48 h of birth, pups were injected intraperitoneally with 50 µl of GBS culture. Challenge inocula were prepared starting from frozen cultures diluted to the appropriate concentration with THB before use. In preliminary experiments (not shown), the challenge doses per pup for each strain tested were determined to cause 90% lethality. Survival of pups was monitored for 2 days after challenge. Protection was calculated as (percentage

$$\frac{\text{percentage deadVaccine} - \text{percentage deadControl}}{\text{percentage deadControl}} \times 100$$

100. Data were evaluated for statistical significance by Fisher's exact test.

Embodiments of the Invention

The invention encompasses, but is not limited to, the embodiments enumerated below.

1. An immunogenic composition comprising a purified Group B Streptococcus (GBS) adhesin island (AI) polypeptide in oligomeric form.

2. The immunogenic composition of embodiment 1 wherein the GBS AI polypeptide is selected from a GBS AI-1.

3. The immunogenic composition of embodiment 1 wherein the GBS AI polypeptide is selected from a GBS AI-2.

1. An immunogenic composition comprising a purified Group B Streptococcus (GBS) adhesin island (AI) polypeptide in oligomeric form.

2. The immunogenic composition of embodiment 1 wherein the GBS AI polypeptide is selected from a GBS AI-1.

3. The immunogenic composition of embodiment 1 wherein the GBS AI polypeptide is selected from a GBS AI-2.

4. The immunogenic composition of any of embodiments 1-3 wherein the GBS AI polypeptide comprises a sortase substrate motif.

5. The immunogenic composition of embodiment 4 wherein the sortase substrate motif is an LPXTG motif.

6. The immunogenic composition of embodiment 5 wherein the LPXTG motif is represented by the amino acid sequence XPXTG, wherein the X at amino acid position 1 is an L, an I, or an F and the X at amino acid position 3 is any amino acid residue.

7. The immunogenic composition of any one of embodiments 1-3 wherein the GBS AI polypeptide affects the ability of GBS bacteria to adhere to epithelial cells.

8. The immunogenic composition of any one of embodiments 1-3 wherein the GBS AI polypeptide affects the ability of GBS bacteria to invade epithelial cells.

9. The immunogenic composition of any one of embodiments 1-3 wherein the GBS AI polypeptide affects the ability of GBS bacteria to translocate through an epithelial cell layer.

10. The immunogenic composition of any one of embodiments 1-3 wherein the GBS AI polypeptide is capable of associating with an epithelial cell surface.

11. The immunogenic composition of embodiment 10 wherein the associating with an epithelial cell surface is binding to the epithelial cell surface.

12. The immunogenic composition of any of embodiments 1-3 wherein the GBS AI polypeptide is a full-length GBS AI protein.

13. The immunogenic composition of any of embodiments 1-3 wherein the GBS AI polypeptide is a fragment of a full-length GBS AI protein.

14. The immunogenic composition of embodiment 13 wherein the fragment comprises at least 7 contiguous amino acid residues of the GBS AI protein.

15. The immunogenic composition of embodiment 2 wherein the GBS AI polypeptide is selected from the group consisting of GBS 80, GBS 104, GBS 52, and fragments thereof.

16. The immunogenic composition of embodiment 3 wherein the GBS AI polypeptide is selected from the group consisting of GBS 59, GBS 67, GBS 150, 01521, 01523, 01524, and fragments thereof.

17. The immunogenic composition of embodiment 15 wherein the GBS AI polypeptide is GBS 80.

18. The immunogenic composition of any of embodiments 1-3 or 15-17 wherein the oligomeric form is a hyperoligomer.

19. The immunogenic composition of any of embodiments 1-3, or 15-17 further comprising a Gram positive bacterium antigen not associated with an AI.

20. The immunogenic composition of embodiment 19 wherein the antigen is selected from the group consisting of GBS 322 and GBS 276.

21. The immunogenic composition of embodiment 20 wherein the antigen is GBS 322.

22. An immunogenic composition comprising a purified Gram positive bacteria adhesin island (AI) polypeptide in an oligomeric form.

23. The immunogenic composition of embodiment 22 wherein the Gram positive bacteria is of a genus selected from the group consisting of *Streptococcus*, *Enterococcus*, *Staphylococcus*, or *Listeria*.

24. The immunogenic composition of embodiment 23 wherein the Gram positive bacteria is of the genus *Streptococcus*.

25. The immunogenic composition of any of embodiments 22-24 wherein the Gram positive bacteria AI polypeptide comprises a sortase substrate motif.

26. The immunogenic composition of embodiment 25 wherein the sortase substrate motif is an LPXTG motif.

27. The immunogenic composition of any one of embodiments 22-24 wherein the Gram positive bacteria AI polypeptide affects the ability of Gram positive bacteria to adhere to epithelial cells.

28. The immunogenic composition of any one of embodiments 22-24 wherein the Gram positive bacteria AI polypeptide affects the ability of Gram positive bacteria to invade epithelial cells.

29. The immunogenic composition of any one of embodiments 22-24 wherein the Gram positive bacteria AI polypeptide affects the ability of Gram positive bacteria to translocate through an epithelial cell layer.

30. The immunogenic composition of any one of embodiments 22-24 wherein the Gram positive bacteria AI polypeptide is capable of associating with an epithelial cell surface.

31. The immunogenic composition of embodiment 30 wherein the associating with an epithelial cell surface is binding to the epithelial cell surface.

32. The immunogenic composition of any of embodiments 22-24 wherein the Gram positive bacteria AI polypeptide is a full-length Gram positive bacteria AI protein.

33. The immunogenic composition of any of embodiments 22-24 wherein the Gram positive bacteria AI polypeptide is a fragment of a full-length Gram positive bacteria AI protein.

34. The immunogenic composition of embodiment 33 wherein the fragment comprises at least 7 contiguous amino acid residues of the Gram positive bacteria AI protein.

35. The immunogenic composition of embodiment 24 wherein the genus *Streptococcus* bacteria is Group A *Streptococcus* (GAS) bacteria and the Gram positive bacteria AI polypeptide is a GAS AI polypeptide.

36. The immunogenic composition of embodiment 35 wherein the GAS AI polypeptide is selected from a GAS AI-1.

37. The immunogenic composition of embodiment 35 wherein the GAS AI polypeptide is selected from a GAS AI-2.

38. The immunogenic composition of embodiment 35 wherein the GAS AI polypeptide is selected from a GAS AI-3.

39. The immunogenic composition of embodiment 35 wherein the GAS AI polypeptide is selected from a GAS AI-4.

40. The immunogenic composition of any of embodiments 35-39 wherein the GAS AI polypeptide comprises a sortase substrate motif.

41. The immunogenic composition of embodiment 40 wherein the sortase substrate motif is an LPXTG motif.

42. The immunogenic composition of embodiment 41 wherein the LPXTG motif is represented by XXXXG, wherein the X at the first amino acid position is an L, a V, an E, or a Q, wherein the X at the second amino acid position is P if the X at the first amino acid position is an L, the X at the second amino acid position is a V if the X at the first amino acid position is an E or a Q, or the X at the second amino acid position is a V or a P if the X at the first amino acid position is a V, wherein the X at the third amino acid position is any amino acid residue, and wherein the X at the fourth amino acid position is a T if the X at the first amino acid position is a V, an E, or a Q, or the X at the fourth amino acid position is a T, an S, or an A if the X at the first amino acid position is an L.

43. The immunogenic composition of any one of embodiments 35-39 wherein the GAS AI polypeptide affects the ability of GAS bacteria to adhere to epithelial cells.

44. The immunogenic composition of any one of embodiments 35-39 wherein the GAS AI polypeptide affects the ability of GAS bacteria to invade epithelial cells.

45. The immunogenic composition of any one of embodiments 35-39 wherein the GAS AI polypeptide affects the ability of GAS bacteria to translocate through an epithelial cell layer.

46. The immunogenic composition of any one of embodiments 35-39 wherein the GAS AI polypeptide is capable of associating with an epithelial cell surface.

47. The immunogenic composition of embodiment 46 wherein the associating with an epithelial cell surface is binding to the epithelial cell surface.

48. The immunogenic composition of any of embodiments 35-39 wherein the GAS AI polypeptide is a full-length GAS AI protein.

49. The immunogenic composition of any of embodiments 35-39 wherein the GAS AI polypeptide is a fragment of a full-length GAS AI protein.

50. The immunogenic composition of embodiment 49 wherein the fragment comprises at least 7 contiguous amino acid residues of the GAS AI protein.

51. The immunogenic composition of embodiment 36 wherein the GAS AI-1 polypeptide is selected from the group consisting of M6_Spy0157, M6_Spy0159, M6_Spy0160, CDC SS 410_fimbrial, ISS3650_fimbrial, DSM2071_fimbrial, and fragments thereof.

52. The immunogenic composition of embodiment 37 wherein the GAS AI-2 polypeptide is selected from the group consisting of GAS15, GAS16, GAS18, and fragments thereof.

53. The immunogenic composition of embodiment 38 wherein the GAS AI-3 polypeptide is selected from the group consisting of SpyM3_0098, SpyM3_0100, SpyM3_0102, SpyM3_0104, SPs0100, SPs0102, SPs0104, SPs0106, orf78, orf80, orf82, orf84, spyM18_0126, spyM18_0128, spyM18_0130, spyM18_0132, SpyoM01000156, SpyoM01000155, SpyoM01000154, SpyoM01000153, SpyoM01000152, SpyoM01000151, SpyoM01000150, SpyoM01000149, ISS3040_fimbrial, ISS3776_fimbrial, ISS4959_fimbrial, and fragments thereof.

53. The immunogenic composition of embodiment 39 wherein the GAS AI-4 polypeptide is selected from the group consisting of 19224134, 19224135, 19224137, 19224139, 19224141, 20010296_fimbrial, 20020069_fimbrial, CDC SS 635_fimbrial, ISS4883_fimbrial, ISS4538_fimbrial, and fragments thereof.

54. The immunogenic composition of embodiment 24 wherein the *Streptococcus* bacteria is *Streptococcus pneumoniae* and the Gram positive bacteria AI polypeptide is a *S. pneumoniae* AI polypeptide.

55. The immunogenic composition of embodiment 54 wherein the *S. pneumoniae* AI polypeptide comprises a sortase substrate motif.

56. The immunogenic composition of embodiment 55 wherein the sortase substrate motif is an LPXTG motif.

57. The immunogenic composition of embodiment 54 wherein the *S. pneumoniae* AI polypeptide affects the ability of *S. pneumoniae* to adhere to epithelial cells.

58. The immunogenic composition of embodiment 54 wherein the *S. pneumoniae* AI polypeptide affects the ability of *S. pneumoniae* to invade epithelial cells.

59. The immunogenic composition of embodiment 54 wherein the *S. pneumoniae* AI polypeptide affects the ability of *S. pneumoniae* to translocate through an epithelial cell layer.

60. The immunogenic composition of embodiment 54 wherein the *S. pneumoniae* AI polypeptide is capable of associating with an epithelial cell surface.

61. The immunogenic composition of embodiment 60 wherein the associating with an epithelial cell surface is binding to the epithelial cell surface.

62. The immunogenic composition of embodiment 54 wherein the *S. pneumoniae* AI polypeptide is a full-length *S. pneumoniae* AI protein.

63. The immunogenic composition of embodiment 54 wherein the *S. pneumoniae* AI polypeptide is a fragment of a full-length *S. pneumoniae* AI protein.

64. The immunogenic composition of embodiment 63 wherein the fragment comprises at least 7 contiguous amino acid residues of the *S. pneumoniae* AI protein.

65. The immunogenic composition of embodiment 54 wherein the *S. pneumoniae* AI polypeptide is selected from the group consisting of SP0462, SP0463, SP0464, orf3_670, orf4_670, orf5_670, ORF3_14CSR, ORF4_14CSR, ORF5_14CSR, ORF3_19AH, ORF4_19AH, ORF5_19AH, ORF3_19FTW, ORF4_19FTW, ORF5_19FTW, ORF3_23FP, ORF4_23FP, ORF5_23FP, ORF3_23FTW, ORF4_23FTW, ORF5_23FTW, ORF3_6BF, ORF4_6BF, ORF5_6BF, ORF3_6BSP, ORF4_6BSP, ORF5_6BSP, ORF3_9VSP, ORF4_9VSP, ORF5_9VSP, and fragments thereof.

66. The immunogenic composition of any one of embodiments 22-24, 35-39, 51-54, or 65 wherein the oligomeric form is a hyperoligomer.

67. The immunogenic composition of any one of embodiments 22-24, 35-39, 51-54, or 65 further comprising a Gram positive bacteria antigen not associated with an AI.

68. The immunogenic composition of embodiment 67 wherein the antigen is selected from the group consisting of GBS 322 and GBS 276.

69. An immunogenic composition comprising a first and a second Group B Streptococcus (GBS) adhesin island (AI) polypeptide.

70. The immunogenic composition of embodiment 69 wherein a full-length polynucleotide sequence encoding for the first GBS AI polypeptide is not present in a GBS bacteria genome comprising a polynucleotide sequence encoding for the second GBS AI polypeptide.

71. The immunogenic composition of embodiment 69 wherein polynucleotides encoding the first and the second GBS AI polypeptide are each present in genomes of more than one GBS serotype and strain isolate.

72. The immunogenic composition of embodiment 69 wherein the first GBS AI polypeptide is encoded by a GBS AI-1.

73. The immunogenic composition of embodiment 69 wherein the first GBS AI polypeptide is encoded by a GBS AI-2.

74. The immunogenic composition of embodiment 72 wherein the second GBS AI polypeptide is encoded by a GBS AI-2.

75. The immunogenic composition of embodiment 73 wherein the second GBS AI polypeptide is encoded by a GBS AI-2.

76. The immunogenic composition of embodiment 72 wherein the second GBS AI polypeptide is encoded by a GBS AI-1.

66 77. The immunogenic composition of embodiment 73 wherein the second GBS AI polypeptide is encoded by a GBS AI-1.

78. The immunogenic composition of embodiment 72 wherein the first GBS AI polypeptide is selected from the group consisting of GBS 80, GBS 104, GBS 52, and fragments thereof.

5 79. The immunogenic composition of embodiment 73 wherein the first GBS AI polypeptide is selected from the group consisting of GBS 59, GBS 67, GBS 150, 01521, 01523, 01524, and fragments thereof.

80. The immunogenic composition of embodiment 74 or 75 wherein the second GBS AI polypeptide is selected from the group consisting of GBS 59, GBS 67, GBS 150, 01521, 01523, 10 01524, and fragments thereof, and wherein the first and the second GBS AI polypeptide are not the same polypeptide.

81. The immunogenic composition of embodiment 76 or 77 wherein the second GBS AI polypeptide is selected from the group consisting of GBS 80, GBS 104, GBS 52, and fragments thereof, and wherein the first and the second GBS AI polypeptide are not the same polypeptide.

15 82. The immunogenic composition of any one of embodiments 69-77 wherein the first GBS AI polypeptide comprises a sortase substrate motif.

83. The immunogenic composition of embodiment 82 wherein the sortase substrate motif is an LPXTG motif.

84. The immunogenic composition of embodiment 83 wherein the LPXTG motif is represented by the sequence XPXTG, wherein the X at amino acid position 1 is an L, an I, or an F and the X at amino acid position 3 is any amino acid residue.

85. The immunogenic composition of any one of embodiments 69-77 wherein the first GBS AI polypeptide affects the ability of GBS bacteria to adhere to epithelial cells.

86. The immunogenic composition of any one of embodiments 69-77 wherein the first GBS 25 AI polypeptide affects the ability of GBS bacteria to invade epithelial cells.

87. The immunogenic composition of any one of embodiments 69-77 wherein the first GBS AI polypeptide affects the ability of GBS bacteria to translocate through an epithelial cell layer.

88. The immunogenic composition of any one of embodiments 69-77 wherein the first GBS AI polypeptide is capable of associating with an epithelial cell surface.

30 89. The immunogenic composition of embodiment 88 wherein the associating with an epithelial cell surface is binding to the epithelial cell surface.

90. The immunogenic composition of any of embodiments 69-77 wherein the first GBS AI polypeptide is a full-length GBS AI protein.

91. The immunogenic composition of any of embodiments 69-77 wherein the first GBS AI 35 polypeptide is a fragment of a full-length GBS AI protein.

92. The immunogenic composition of embodiment 91 wherein the fragment comprises at least 7 contiguous amino acid residues of the first GBS AI protein.

93. The immunogenic composition of any one of embodiments 69-79 wherein the first GBS AI polypeptide is in oligomeric form.

94. The immunogenic composition of any one of embodiments 69-77 wherein the second GBS AI polypeptide is in oligomeric form.

95. The immunogenic composition of any one of embodiments 69-79 wherein the first and the second GBS AI polypeptide are associated in a single oligomeric form.

96. The immunogenic composition of embodiment 95 wherein the first and the second GBS AI polypeptides are chemically associated.

97. The immunogenic composition of embodiment 95 wherein the first and the second GBS AI polypeptides are physically associated.

98. The immunogenic composition of embodiment 93 wherein the oligomeric form is a hyperoligomer.

99. The immunogenic composition of embodiment 94 wherein the oligomeric form is a hyperoligomer.

100. The immunogenic composition of embodiment 76 wherein the first GBS AI polypeptide is GBS 80 and the second GBS AI polypeptide is GBS 104.

101. The immunogenic composition of embodiment 74 wherein the first GBS AI polypeptide is GBS 80 and the second GBS AI polypeptide is GBS 67.

102. The immunogenic composition of any one of embodiments 69-79, 100, or 101 further comprising a GBS polypeptide not associated with an AI.

103. The immunogenic composition of embodiment 102 wherein the GBS polypeptide not associated with an AI is selected from the group consisting of GBS 322 and GBS 276.

104. The immunogenic composition of embodiment 103 wherein the GBS polypeptide not associated with an AI is GBS 322.

105. An immunogenic composition comprising a first and a second Gram positive bacteria adhesin island (AI) polypeptide.

106. The immunogenic composition of embodiment 105 wherein a full length polynucleotide sequence encoding for the first Gram positive bacteria AI polypeptide is not present in a genome of a Gram positive bacteria comprising a full length polynucleotide sequence encoding for the second Gram positive bacteria AI polypeptide.

107. The immunogenic composition of embodiment 105 wherein polynucleotides encoding the first and the second Gram positive bacteria AI polypeptide are each present in genomes of more than one Gram positive bacteria serotype and strain isolate.

108. The immunogenic composition of embodiment 105 wherein the first and the second Gram positive bacteria AI polypeptides are of different Gram positive bacteria species.

109. The immunogenic composition of embodiment 105 wherein the first and the second Gram positive bacteria AI polypeptides are of the same Gram positive bacteria species.

110. The immunogenic composition of embodiment 105 wherein the first and the second Gram positive bacteria AI polypeptides are from different AI subtypes.

111. The immunogenic composition of embodiment 105 wherein the first and the second Gram positive bacteria AI polypeptides are from the same AI subtype.

112. The immunogenic composition of embodiment 105 wherein the first Gram positive bacteria AI polypeptide has detectable surface exposure on a first Gram positive bacteria strain or serotype but not a second Gram positive bacteria strain or subtype and the second Gram positive bacteria AI polypeptide has detectable surface exposure on the second Gram positive bacteria strain or serotype but not the first Gram positive bacteria strain or serotype.

113. The immunogenic composition of embodiment 105 wherein the Gram positive bacteria is *S. pneumoniae*, *S. mutans*, *E. faecalis*, *E. faecium*, *C. difficile*, *L. monocytogenes*, or *C. diphtheriae*.

114. The immunogenic composition of any of embodiments 105-113 wherein the first and the second Gram positive bacteria AI polypeptides comprise a sortase substrate motif.

115. The immunogenic composition of embodiment 114 wherein the sortase substrate motif is an LPXTG motif.

116. The immunogenic composition of embodiment 115 wherein the LPXTG motif is represented by XXXXG, wherein the X at amino acid position 1 is an L, a V, an E, an I, an F, or a Q, wherein X at amino acid position 2 is a P if X at amino acid position 1 is an L, an I, or an F, wherein X at amino acid position 2 is a V if X at amino acid position 1 is a E or a Q, wherein X at amino acid position 2 is a V or a P if X at amino acid position 1 is a V, wherein X at amino acid position 3 is any amino acid residue, wherein X at amino acid position 4 is a T if X at amino acid position 1 is a V, E, I, F, or Q, and wherein X at amino acid position 4 is a T, S, or A if X at amino acid position 1 is an L.

117. The immunogenic composition of embodiment 105 wherein the first Gram positive bacteria AI polypeptide is a first Group A Streptococcus (GAS) AI polypeptide.

118. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide comprises a sortase substrate motif.

119. The immunogenic composition of embodiment 118 wherein the sortase substrate motif is an LPXTG motif.

120. The immunogenic composition of embodiment 119 wherein the LPXTG motif is represented by XXXXG, wherein the X at the first amino acid position is an L, a V, an E, or a Q, wherein the X at the second amino acid position is P if the X at the first amino acid position is an L, the X at the second amino acid position is a V if the X at the first amino acid position is an E or a Q, or the X at the second amino acid position is a V or a P if the X at the first amino acid position is a V, wherein the X at the third amino acid position is any amino acid residue, and wherein the X at the fourth amino acid position is a T if the X at the first amino acid position is a V, an E, or a Q, or the X at the fourth amino acid position is a T, an S, or an A if the X at the first amino acid position is an L.

121. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide affects the ability of GAS bacteria to adhere to epithelial cells.

122. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide affects the ability of GAS bacteria to invade epithelial cells.

123. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide affects the ability of GAS bacteria to translocate through an epithelial cell layer.

124. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide is capable of associating with an epithelial cell surface.

125. The immunogenic composition of embodiment 117 wherein the associating with an epithelial cell surface is binding to the epithelial cell surface.

126. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide is a full-length GAS AI protein.

127. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide is a fragment of a full-length GAS AI protein.

128. The immunogenic composition of embodiment 127 wherein the fragment comprises at least 7 contiguous amino acid residues of the GAS AI protein.

129. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide is a first GAS AI-1 polypeptide.

130. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide is a first GAS AI-2 polypeptide.

131. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide is a first GAS AI-3 polypeptide.

132. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide is a first GAS AI-4 polypeptide.

133. The immunogenic composition of any one of embodiments 117 or 129-132 wherein the second Gram positive bacteria AI polypeptide is a second GAS AI polypeptide.

134. The immunogenic composition of embodiment 133 wherein the second GAS AI polypeptide is a second GAS AI-1 polypeptide.

135. The immunogenic composition of embodiment 133 wherein the second GAS AI polypeptide is a second GAS AI-2 polypeptide.

136. The immunogenic composition of embodiment 133 wherein the second GAS AI polypeptide is a second GAS AI-3 polypeptide.

137. The immunogenic composition of embodiment 133 wherein the second GAS AI polypeptide is a second GAS AI-4 polypeptide.

138. The immunogenic composition of embodiment 129 wherein the first GAS AI-1 polypeptide is selected from the group consisting of M6_Spy0157, M6_Spy0159, M6_Spy0160, CDC SS 410_fimbrial, ISS3650_fimbrial, DSM2071_fimbrial, and fragments thereof.

139. The immunogenic composition of embodiment 130 wherein the first GAS AI-2 polypeptide is selected from the group consisting of GAS15, GAS16, GAS18, and fragments thereof.

140. The immunogenic composition of embodiment 131 wherein the first GAS AI-3 polypeptide is selected from the group consisting of SpyM3_0098, SpyM3_0100, SpyM3_0102, SpyM3_0104, SPs0100, SPs0102, SPs0104, SPs0106, orf78, orf80, orf82, orf84, spyM18_0126, spyM18_0128, spyM18_0130, spyM18_0132, SpyoM01000156, SpyoM01000155, SpyoM01000154, SpyoM01000153, SpyoM01000152, SpyoM01000151, SpyoM01000150, SpyoM01000149, ISS3040_fimbrial, ISS3776_fimbrial, ISS4959_fimbrial, and fragments thereof.

141. The immunogenic composition of embodiment 132 wherein the first GAS AI-4 polypeptide is selected from the group consisting of 19224134, 19224135, 19224137, 19224139, 19224141, 20010296_fimbrial, 20020069_fimbrial, CDC SS 635_fimbrial, ISS4883_fimbrial, ISS4538_fimbrial, and fragments thereof.

142. The immunogenic composition of embodiment 134 wherein the second GAS AI-1 polypeptide is selected from the group consisting of M6_Spy0157, M6_Spy0159, M6_Spy0160, CDC SS 410_fimbrial, ISS3650_fimbrial, DSM2071_fimbrial, and fragments thereof.

143. The immunogenic composition of embodiment 135 wherein the second GAS AI-2 polypeptide is selected from the group consisting of GAS15, GAS16, GAS18, and fragments thereof.

144. The immunogenic composition of embodiment 136 wherein the second GAS AI-3 polypeptide is selected from the group consisting of SpyM3_0098, SpyM3_0100, SpyM3_0102, SpyM3_0104, SPs0100, SPs0102, SPs0104, SPs0106, orf78, orf80, orf82, orf84, spyM18_0126, spyM18_0128, spyM18_0130, spyM18_0132, SpyoM01000156, SpyoM01000155, SpyoM01000154, SpyoM01000153, SpyoM01000152, SpyoM01000151, SpyoM01000150, SpyoM01000149, ISS3040_fimbrial, ISS3776_fimbrial, ISS4959_fimbrial, and fragments thereof.

145. The immunogenic composition of embodiment 137 wherein the second GAS AI-4 polypeptide is selected from the group consisting of 19224134, 19224135, 19224137, 19224139, 19224141, 20010296_fimbrial, 20020069_fimbrial, CDC SS 635_fimbrial, ISS4883_fimbrial, ISS4538_fimbrial, and fragments thereof.

146. The immunogenic composition of any one of embodiments 117-132 or 138-141 wherein the second Gram positive bacteria AI polypeptide is a Group B Streptococcus (GBS) AI polypeptide.

147. The immunogenic composition of embodiment 146 wherein the GBS AI polypeptide comprises a sortase substrate motif.

148. The immunogenic composition of embodiment 147 wherein the sortase substrate motif is an LPXTG motif.

149. The immunogenic composition of embodiment 148 wherein the LPXTG motif is represented by the amino acid sequence XPXTG, wherein the X at amino acid position 1 is an L, an I, or an F and the X at amino acid position 3 is any amino acid residue.

150. The immunogenic composition of embodiment 146 wherein the GBS AI polypeptide affects the ability of GBS bacteria to adhere to epithelial cells.

151. The immunogenic composition of embodiment 146 wherein the GBS AI polypeptide affects the ability of GBS bacteria to invade epithelial cells.

152. The immunogenic composition of embodiment 146 wherein the GBS AI polypeptide affects the ability of GBS bacteria to translocate through an epithelial cell layer.

153. The immunogenic composition of embodiment 146 wherein the GBS AI polypeptide is capable of associating with an epithelial cell surface.

154. The immunogenic composition of embodiment 146 wherein the associating with an epithelial cell surface is binding to the epithelial cell surface.

155. The immunogenic composition of embodiment 146 wherein the GBS AI polypeptide is a full-length GBS AI protein.

156. The immunogenic composition of embodiment 146 wherein the GBS AI polypeptide is a fragment of a full-length GBS AI protein.

157. The immunogenic composition of embodiment 156 wherein the fragment comprises at least 7 contiguous amino acid residues of the GBS AI protein.

158. The immunogenic composition of embodiment 146 wherein the GBS AI polypeptide is a GBS AI-1 polypeptide.

159. The immunogenic composition of embodiment 146 wherein the GBS AI polypeptide is a GBS AI-2 polypeptide.

160. The immunogenic composition of embodiment 158 wherein the GBS AI-1 polypeptide is selected from the group consisting of GBS 80, GBS 104, GBS 52, and fragments thereof.

161. The immunogenic composition of embodiment 159 wherein the GBS AI-2 polypeptide is selected from the group consisting of GBS 59, GBS 67, GBS 150, 01521, 01523, 01524, and fragments thereof.

162. The immunogenic composition of any one of embodiments 117-132 or 138-141 wherein the second Gram positive bacteria AI polypeptide is a *Streptococcus pneumoniae* AI polypeptide.

163. The immunogenic composition of embodiment 162 wherein the *S. pneumoniae* AI polypeptide comprises a sortase substrate motif.

164. The immunogenic composition of embodiment 163 wherein the sortase substrate motif is an LPXTG motif.

165. The immunogenic composition of embodiment 162 wherein the *S. pneumoniae* AI polypeptide affects the ability of *S. pneumoniae* to adhere to epithelial cells.

166. The immunogenic composition of embodiment 162 *S. pneumoniae* AI polypeptide affects the ability of *S. pneumoniae* to invade epithelial cells.

167. The immunogenic composition of embodiment 162 wherein the *S. pneumoniae* AI polypeptide affects the ability of *S. pneumoniae* to translocate through an epithelial cell layer.

168. The immunogenic composition of embodiment 162 wherein the *S. pneumoniae* AI polypeptide is capable of associating with an epithelial cell surface.

169. The immunogenic composition of embodiment 168 wherein the associating with an epithelial cell surface is binding to the epithelial cell surface.

170. The immunogenic composition of embodiment 162 wherein the *S. pneumoniae* AI polypeptide is a full-length *S. pneumoniae* AI protein.

171. The immunogenic composition of embodiment 162 wherein the *S. pneumoniae* AI polypeptide is a fragment of a full-length *S. pneumoniae* AI protein.

172. The immunogenic composition of embodiment 162 wherein the fragment comprises at least 7 contiguous amino acid residues of the *S. pneumoniae* AI protein.

173. The immunogenic composition of embodiment 162 wherein the *S. pneumoniae* AI polypeptide is selected from the group consisting of SP0462, SP0463, SP0464, orf3_670, orf4_670, orf5_670, ORF3_14CSR, ORF4_14CSR, ORF5_14CSR, ORF3_19AH, ORF4_19AH, ORF5_19AH, ORF3_19FTW, ORF4_19FTW, ORF5_19FTW, ORF3_23FP, ORF4_23FP, ORF5_23FP, ORF3_23FTW, ORF4_23FTW, ORF5_23FTW, ORF3_6BF, ORF4_6BF, ORF5_6BF, ORF3_6BSP, ORF4_6BSP, ORF5_6BSP, ORF3_9VSP, ORF4_9VSP, ORF5_9VSP, and fragments thereof.

174. The immunogenic composition of any one of embodiments 105-117 wherein the first Gram positive bacteria AI polypeptide is in oligomeric form.

175. The immunogenic composition of embodiment 174 wherein the oligomeric form is a hyperoligomer.

176. The immunogenic composition of embodiment 174 wherein the second Gram positive bacteria AI polypeptide is in oligomeric form.

177. The immunogenic composition of embodiment 176 wherein the oligomeric form is a hyperoligomer.

178. The immunogenic composition of embodiment 176 wherein the first and the second Gram positive bacteria AI polypeptide are associated in a single oligomeric form.

179. The immunogenic composition of embodiment 178 wherein the first and the second Gram positive bacteria AI polypeptide are chemically associated.

180. The immunogenic composition of embodiment 178 wherein the first and the second Gram positive bacteria AI polypeptide are physically associated.

181. The immunogenic composition of any one of embodiments 105-117 further comprising a Gram positive bacteria polypeptide not associated with an AI.

182. The immunogenic composition of embodiment 181 wherein the Gram positive bacteria polypeptide not associated with an AI is selected from the group consisting of GBS 322 and GBS 276.

183. The immunogenic composition of embodiment 182 wherein the Gram positive bacteria polypeptide not associated with an AI is GBS 322.

184. A modified Gram positive bacterium adapted to produce increased levels of AI surface protein.

185. The modified Gram positive bacterium of embodiment 184 wherein the AI surface protein is in oligomeric form.

186. The modified Gram positive bacterium of embodiment 185 wherein the oligomeric form is a hyperoligomer.

187. The modified Gram positive bacterium of any one of embodiments 184-186 which is a Group B Streptococcus bacterium.

188. The modified Gram positive bacterium of any one of embodiments 184-186 which is a Group A Streptococcus bacterium.

189. The modified Gram positive bacterium of any one of embodiments 184-186 which is a non-pathogenic Gram positive bacterium.

190. The modified Gram positive bacterium of embodiment 189 wherein the non-pathogenic Gram positive bacterium is *Streptococcus gordonii*.

191. The modified Gram positive bacterium of embodiment 189 wherein the non-pathogenic Gram positive bacterium is *Lactococcus lactis*.

192. The modified Gram positive bacterium of any one of embodiments 184-186 which has been inactivated and wherein the AI surface protein is exposed on the surface of the Gram positive bacterium.

193. The modified Gram positive bacterium of any one of embodiments 184-186 which has been attenuated and wherein the AI surface protein is exposed on the surface of the Gram positive bacterium.

194. The modified GBS bacterium of embodiment 187 which has been inactivated and wherein the AI surface protein is exposed on the surface of the GBS bacterium.

195. The modified GBS bacterium of embodiment 187 which has been attenuated and wherein the AI surface protein is exposed on the surface of the GBS bacterium.

196. The modified GAS bacterium of embodiment 188 which has been inactivated and wherein the AI surface protein is exposed on the surface of the GAS bacterium.

197. The modified GAS bacterium of embodiment 188 which has been attenuated and wherein the AI surface protein is exposed on the surface of the GAS bacterium.

198. The modified non-pathogenic bacterium of embodiment 189 which has been inactivated and wherein the AI surface protein is exposed on the surface of the non-pathogenic Gram positive bacterium.

199. The modified non-pathogenic bacterium of embodiment 189 which has been attenuated and wherein the AI surface protein is exposed on the surface of the non-pathogenic Gram positive bacterium.

200. A method for manufacturing an oligomeric adhesin island (AI) surface antigen comprising:

culturing a Gram positive bacterium that expresses an oligomeric AI surface antigen and isolating the expressed oligomeric AI surface antigen.

201. The method of embodiment 200 wherein the step of isolating is performed by collecting said oligomeric AI surface antigen from Gram positive bacterium secretions in the Gram positive bacterium culture.

202. The method of embodiment 200 further comprising a step of purifying.

203. The method of embodiment 202 wherein the oligomeric AI surface antigen is purified from the Gram positive bacterium cell surface.

204. The method of embodiment 200 wherein the Gram positive bacterium is adapted for increased AI protein expression.

5 205. The method of any one of embodiments 200-204 wherein the Gram positive bacterium is a Group A Streptococcus bacterium.

206. The method of any one of embodiments 200-204 wherein the Gram positive bacterium is a Group B Streptococcus bacterium.

10 207. The method of any one of embodiments 200-204 wherein the oligomeric AI surface antigen is in hyperoligomeric form.

208. The method of embodiment 200 wherein the Gram positive bacterium expresses the oligomeric AI surface antigen recombinantly.

209. The method of embodiment 208 wherein the Gram positive bacterium further manipulated expresses at least 1 AI sortase.

15 210. The modified Gram positive bacterium of any one of embodiments 184-186 which is a *S. pneumoniae* bacterium.

211. The method of any one of embodiments 200-204 wherein the Gram positive bacterium is *S. pneumoniae*.

1. An immunogenic composition comprising a purified Group B Streptococcus (GBS) adhesin island (AI) polypeptide in oligomeric form.

2. The immunogenic composition of claim 1 wherein the GBS AI polypeptide is selected from a GBS AI-1.

3. The immunogenic composition of claim 1 wherein the GBS AI polypeptide is selected from a GBS AI-2.

4. The immunogenic composition of claim 2 wherein the GBS AI polypeptide is selected from the group consisting of GBS 80, GBS 104, GBS 52, and fragments thereof.

5. The immunogenic composition of claim 3 wherein the GBS AI polypeptide is selected from the group consisting of GBS 59, GBS 67, GBS 150, 01521, 01523, 01524, and fragments thereof.

6. The immunogenic composition of claim 4 wherein the GBS AI polypeptide is GBS 80.

7. The immunogenic composition of any of claims 1-6 wherein the oligomeric form is a hyperoligomer.

8 (22). An immunogenic composition comprising a purified Gram positive bacteria adhesin island (AI) polypeptide in an oligomeric form.

9 (23). The immunogenic composition of claim 8 wherein the Gram positive bacteria is of a genus selected from the group consisting of *Streptococcus*, *Enterococcus*, *Staphylococcus*, *Clostridium*, *Corynebacterium*, or *Listeria*.

10 (24). The immunogenic composition of claim 9 wherein the Gram positive bacteria is of the genus *Streptococcus*.

11 (35). The immunogenic composition of claim 10 wherein the genus *Streptococcus* bacteria is Group A Streptococcus (GAS) bacteria and the Gram positive bacteria AI polypeptide is a GAS AI polypeptide.

12 (36). The immunogenic composition of claim 11 wherein the GAS AI polypeptide is selected from a GAS AI-1.

13 (37). The immunogenic composition of claim 11 wherein the GAS AI polypeptide is selected from a GAS AI-2.

14 (38). The immunogenic composition of claim 11 wherein the GAS AI polypeptide is selected from a GAS AI-3.

15 (39). The immunogenic composition of claim 11 wherein the GAS AI polypeptide is selected from a GAS AI-4.

16 (66). The immunogenic composition of any one of claims 8-15 wherein the oligomeric form is a hyperoligomer.

17. An immunogenic composition comprising a first and a second Group B Streptococcus (GBS) adhesin island (AI) polypeptide.

18. The immunogenic composition of claim 17 wherein the first GBS AI polypeptide is encoded by a GBS AI-1.

19. The immunogenic composition of claim 18 wherein the second GBS AI polypeptide is encoded by a GBS AI-2.

20. The immunogenic composition of claim 18 wherein the first GBS AI polypeptide is selected from the group consisting of GBS 80, GBS 104, GBS 52, and fragments thereof.

21. The immunogenic composition of claim 19 wherein the second GBS AI polypeptide is selected from the group consisting of GBS 59, GBS 67, GBS 150, 01521, 01523, 01524, and fragments thereof, and wherein the first and the second GBS AI polypeptide are not the same polypeptide.

22. The immunogenic composition of claim 19 wherein the first GBS AI polypeptide is GBS 80 and the second GBS AI polypeptide is GBS 67.

23. An immunogenic composition comprising a first and a second Gram positive bacteria adhesin island (AI) polypeptide.

24. The immunogenic composition of claim 23 wherein the Gram positive bacteria is *Streptococcus*, *Enterococcus*, *Staphylococcus*, *Clostridium*, *Corynebacterium*, or *Listeria*.

25. The immunogenic composition of claim 23 wherein the first Gram positive bacteria AI polypeptide is a first Group A *Streptococcus* (GAS) AI polypeptide.

26. The immunogenic composition of claim 25 wherein the first GAS AI polypeptide is a first GAS AI-1 polypeptide.

27. The immunogenic composition of claim 25 wherein the first GAS AI polypeptide is a first GAS AI-2 polypeptide.

28. The immunogenic composition of claim 25 wherein the first GAS AI polypeptide is a first GAS AI-3 polypeptide.

29. The immunogenic composition of claim 25 wherein the first GAS AI polypeptide is a first GAS AI-4 polypeptide.

30. The immunogenic composition of any one of claims 25-29 wherein the second Gram positive bacteria AI polypeptide is a second GAS AI polypeptide.

31. The immunogenic composition of claim 30 wherein the second GAS AI polypeptide is a second GAS AI-1 polypeptide.

32. The immunogenic composition of claim 30 wherein the second GAS AI polypeptide is a second GAS AI-2 polypeptide.

33. The immunogenic composition of claim 30 wherein the second GAS AI polypeptide is a second GAS AI-3 polypeptide.

34. The immunogenic composition of claim 30 wherein the second GAS AI polypeptide is a second GAS AI-4 polypeptide.

35. A modified Gram positive bacterium adapted to produce increased levels of AI surface protein.

36. The modified Gram positive bacterium of claim 35 wherein the AI surface protein is in oligomeric form.

37. The modified Gram positive bacterium of claim 36 wherein the oligomeric form is a hyperoligomer.

5 38. The modified Gram positive bacterium of any one of claims 35-37 which is a non-pathogenic Gram positive bacterium.

39. The modified Gram positive bacterium of claim 38 wherein the non-pathogenic Gram positive bacterium is *Lactococcus lactis*.

10 40. A method for manufacturing an oligomeric adhesin island (AI) surface antigen comprising:

culturing a Gram positive bacterium that expresses an oligomeric AI surface antigen and isolating the expressed oligomeric AI surface antigen.

PCT/US2005/027239

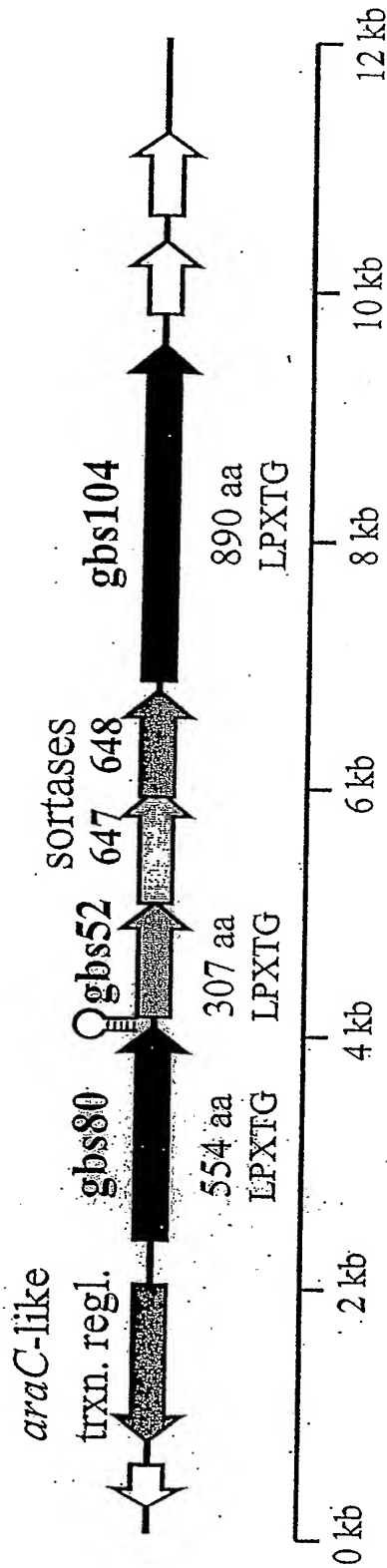
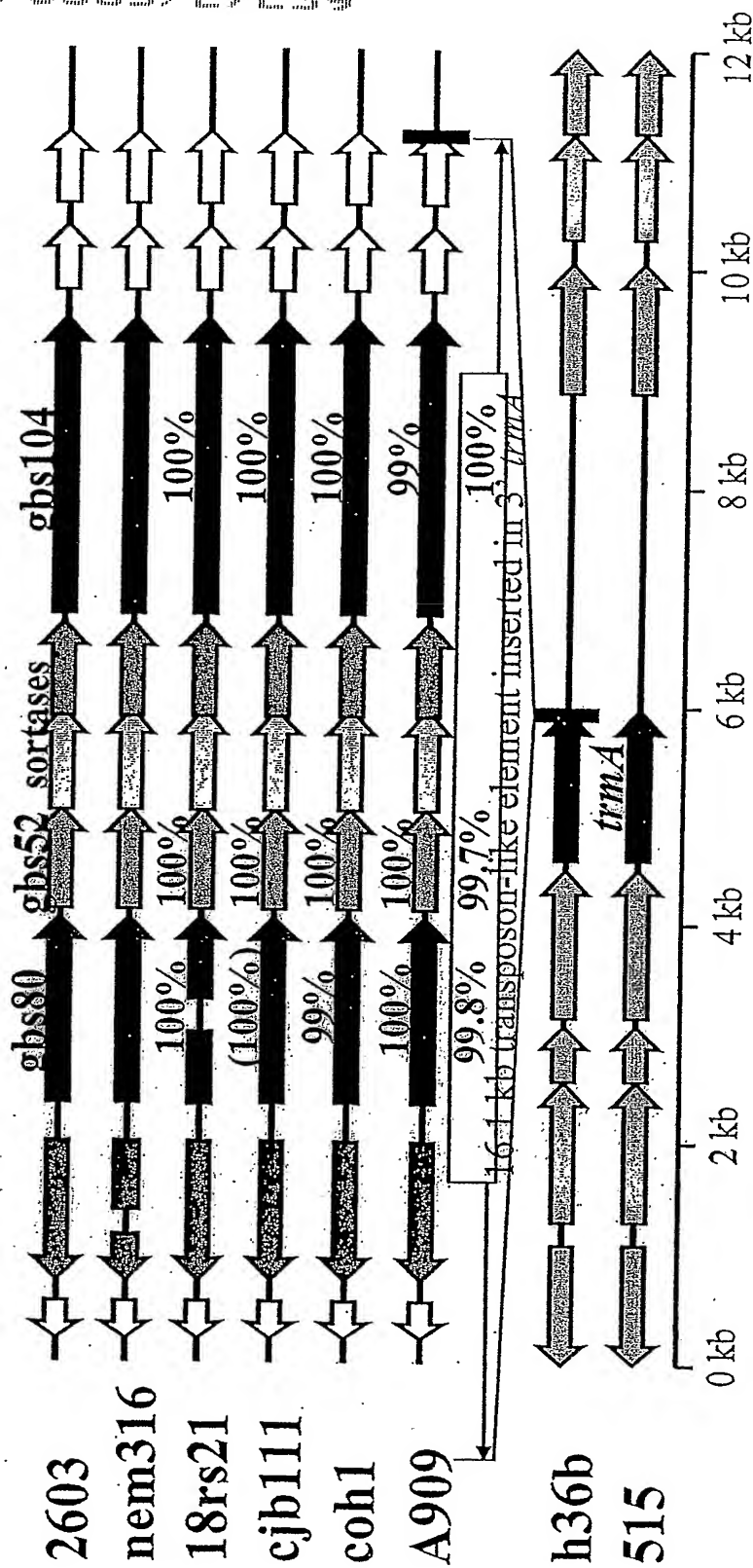
FIGURE 1: Adhesion Island 1

Figure 2: Conservation of AI-1 in GBS serotypes and strain isolates



PCT/US2005/027239

FIGURE 3: Correlation of AI-1 and AI-2 within GBS serotype V, strain isolate 2603 genome

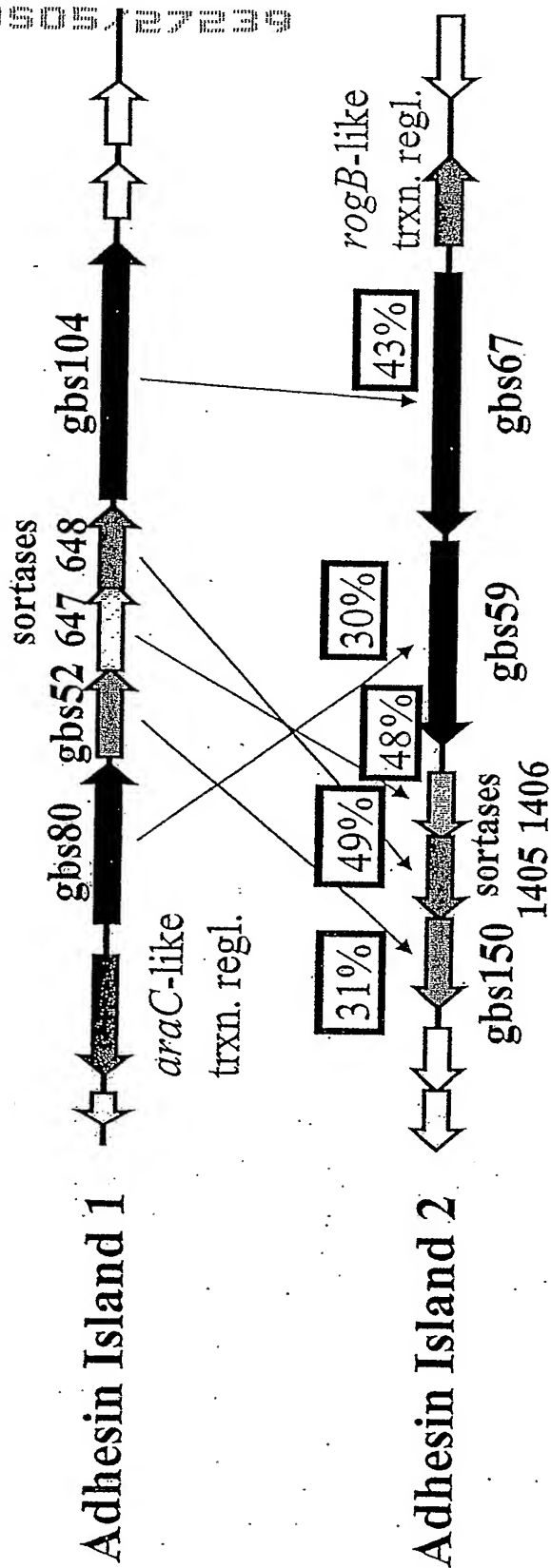
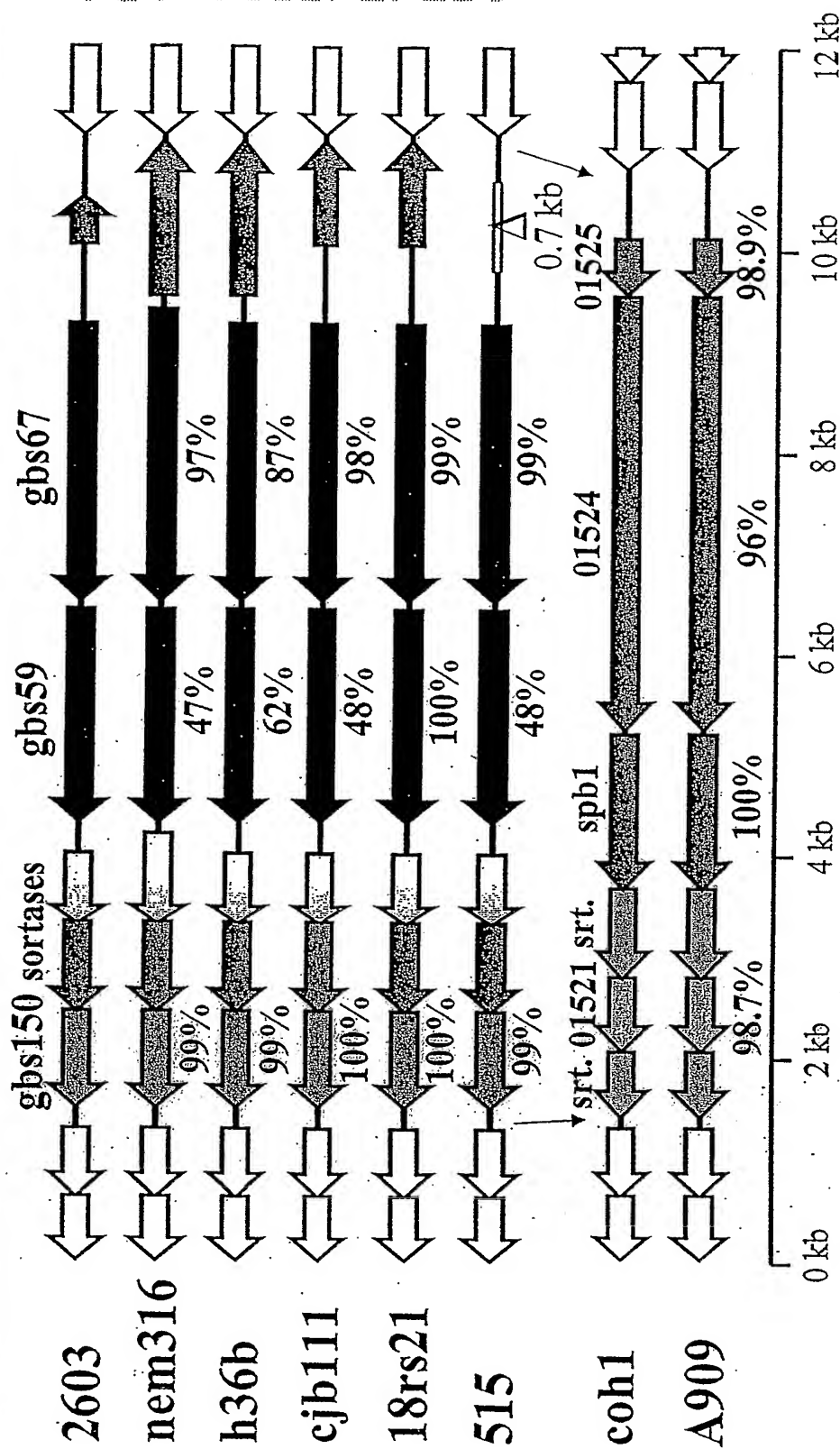


Figure 4: Identification and Variance of AI-2 in Several GBS Serotypes and Strain Isolates



A:

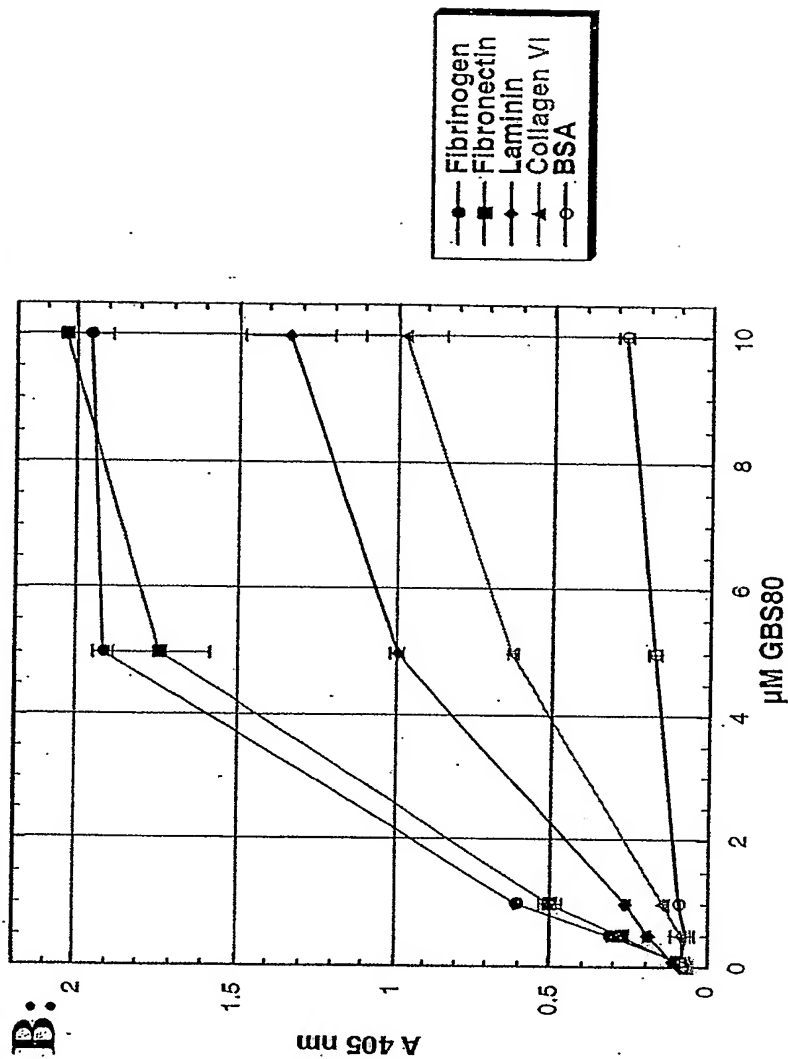
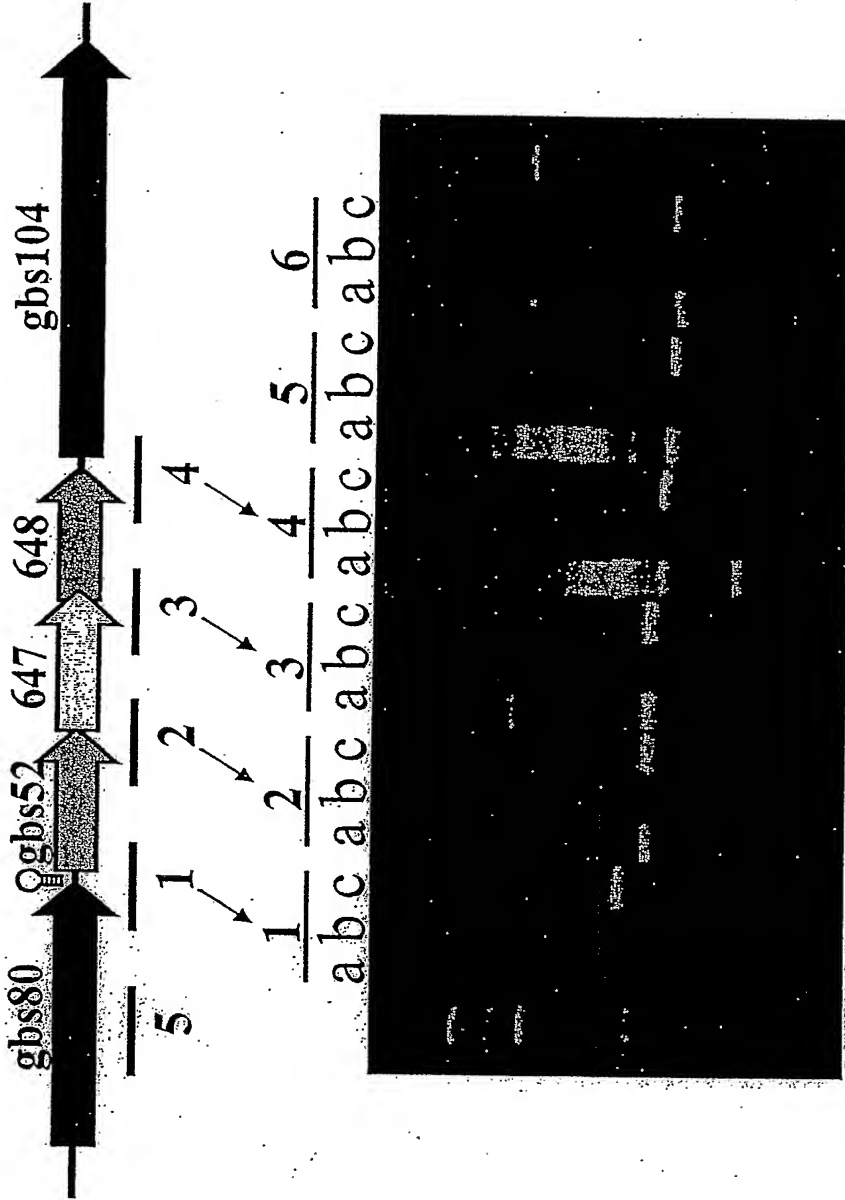


Figure 6: Adhesion Island I is an operon by RT-PCR



7/487

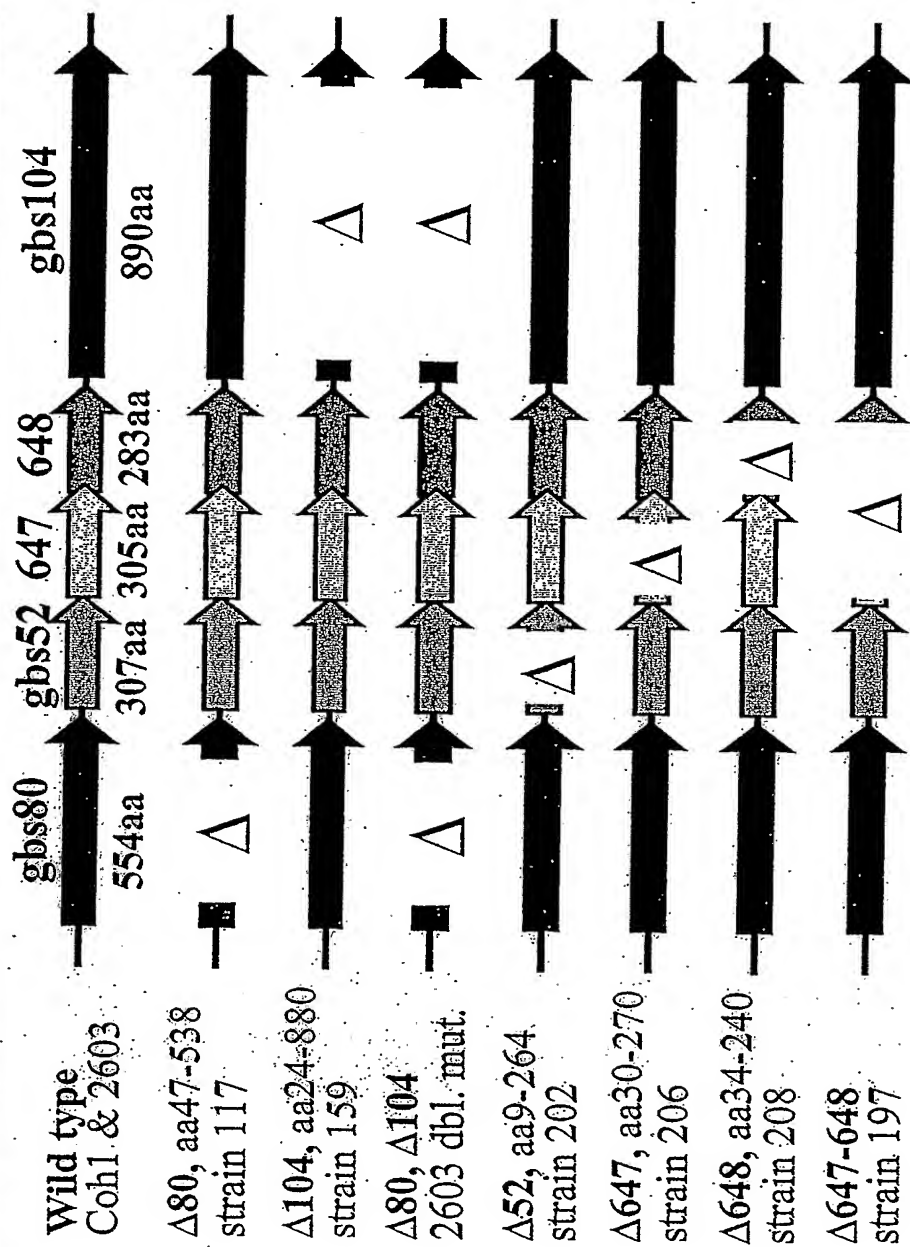
Figure 7: In frame deletions of AI-1 genes

Figure 8: gbs80 is required for surface localization of gbs104

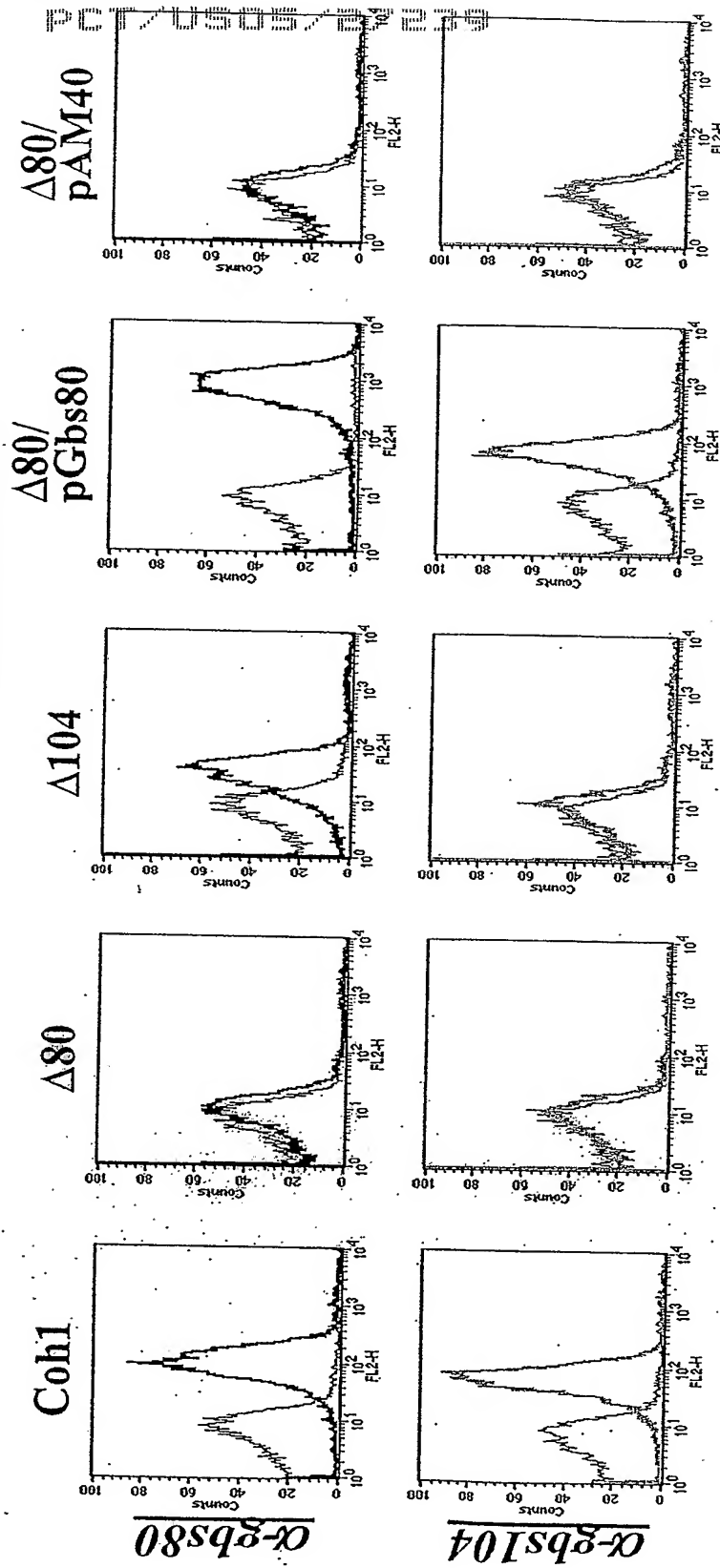


Figure 9: sortases 647 & 648 play a semi-redundant role in surface exposure of gbs80 and gbs104

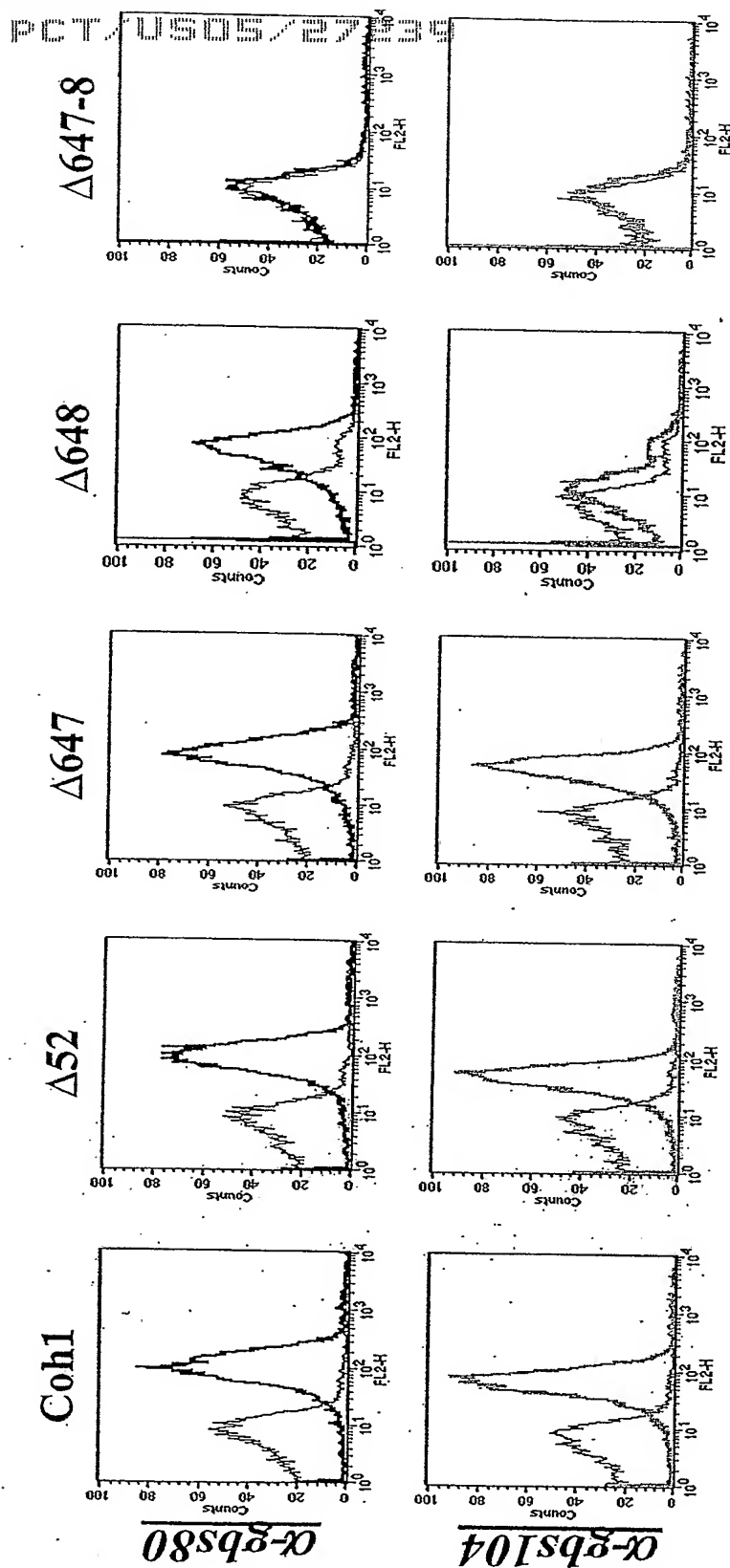
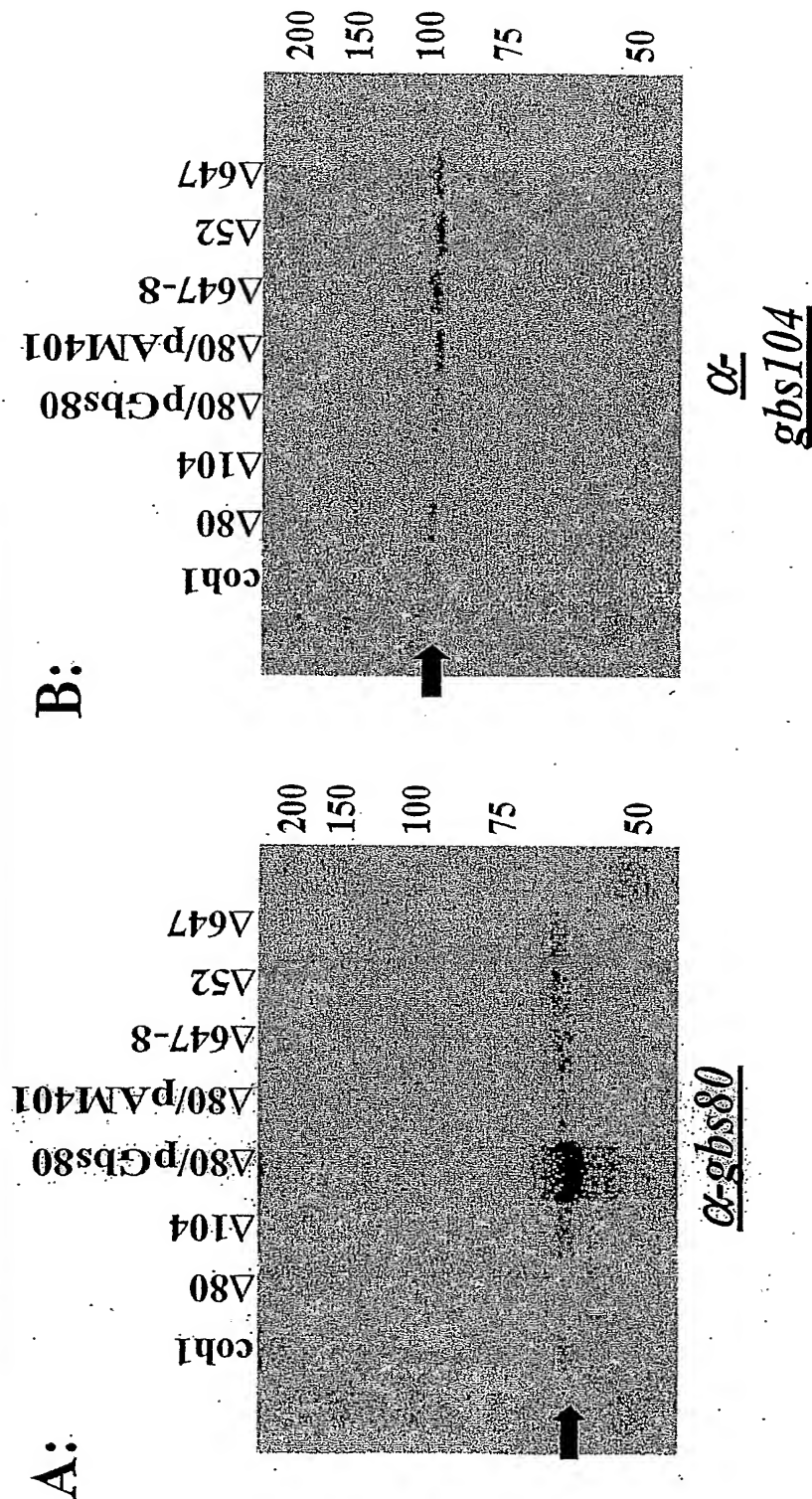
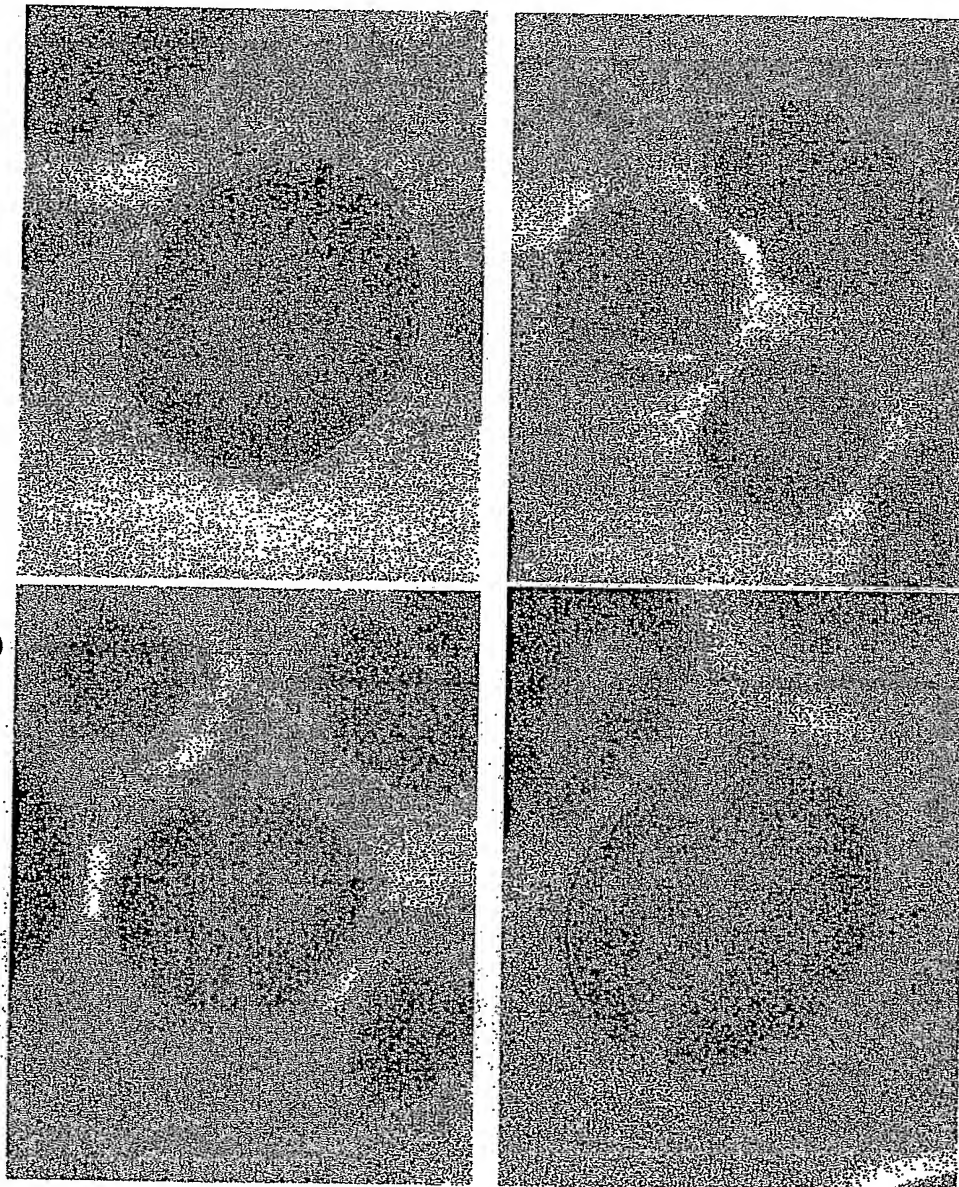


Figure 10: Western blot of mutant strains



**FIGURE 11: Pre-embedding IEM
staining of GBS 80**



PCT/US05/27239 12/487

FIGURE 12: Predicted Secondary Structure for GBS 067

PHD SECONDARY STRUCTURE PREDICTION for GBS 067

```

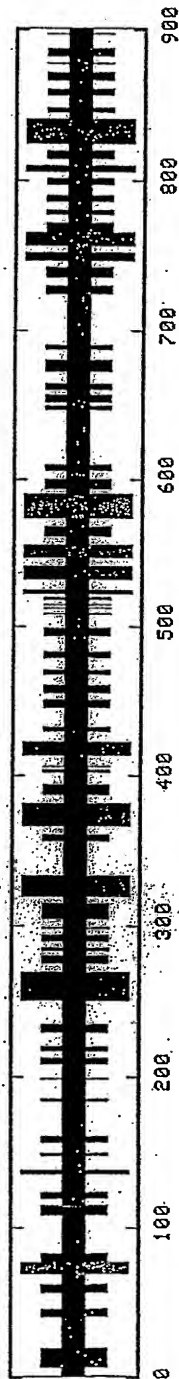
10 20 30 40 50 60 70
MRKYQFSKILTSLFUISQIPLNTNVLGESTVPEAGKGLVVKRTDDQNKPLSKATFVLKTTAHPESK
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
TEKVAELTGEATFDNLIPGDYTLSEETAPEGYKTNQTWQVKSNGKTTIIONSGDKNSTIGONQBEELD
HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
KQYPTGIYEDTKESYLEHVKGSPNGKSEAKVNPYSSEGEHIREIPGTLSKRISSEVGLAHNKYKI
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
ELTVSGKTIKVPDKQKPLDVFLVLDNSNSMNDGNFQRHNKAKAABALGTAVDILGANSNDNRVALV
EEEECCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
TYGSDIFDGRSDVVKGFKDDKYGLQTKFTIGTENYSHKQLTNABEIIKRIPTAPKAKWGSTTNGL
ECCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
TPEQKEYYLSKVGETFTWKAEMEADDILSQVNRANSOKIIVHTDGPVTRSYAINNFKLGASYESQFEOM
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
KNGYLNKSNFLLPDKPEDIKNGESYFLPLDSYOTQIISGLQLKHYLDLNLNYPKGTIYRNGPVKEH
HHHhCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
GTPTKLYINSLKQKNYDIFNFGIDISGFRQVNHVYKKNQDGTQKLEAEAFKLSDGITELMRFSKSP
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
EYTPIVTSADTSNNHILSKIQQOFETILFKNSLVANGTIEDPMGDKINLQNGGOTLPQSDYTLQNDG
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
SVMKDGIATGPNNDGILKGVLEYIGNKLYVRGINLGEQKVLTLYDVKLDSDFSINKFYDNGRRTL
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
NPKSDDPNTLRDFFPKIRDVREYPTITIKNEKNGEIEFIKVDKDNKKLLKGAFFELQEFNEDYKLYL
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
PIKNNNSKVVTGENGISYKDLKGYQLIEAVSPEDVQKINPKPILTFEVKGSIKNIIAVNKQISEYH
eCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
EEGDKHLITNTHIPPKGILPMTGGKILSPILIGGAMSIAGGIYTWKRYKKSDDMSIKKD
HhCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC

```

Sequence length: 901

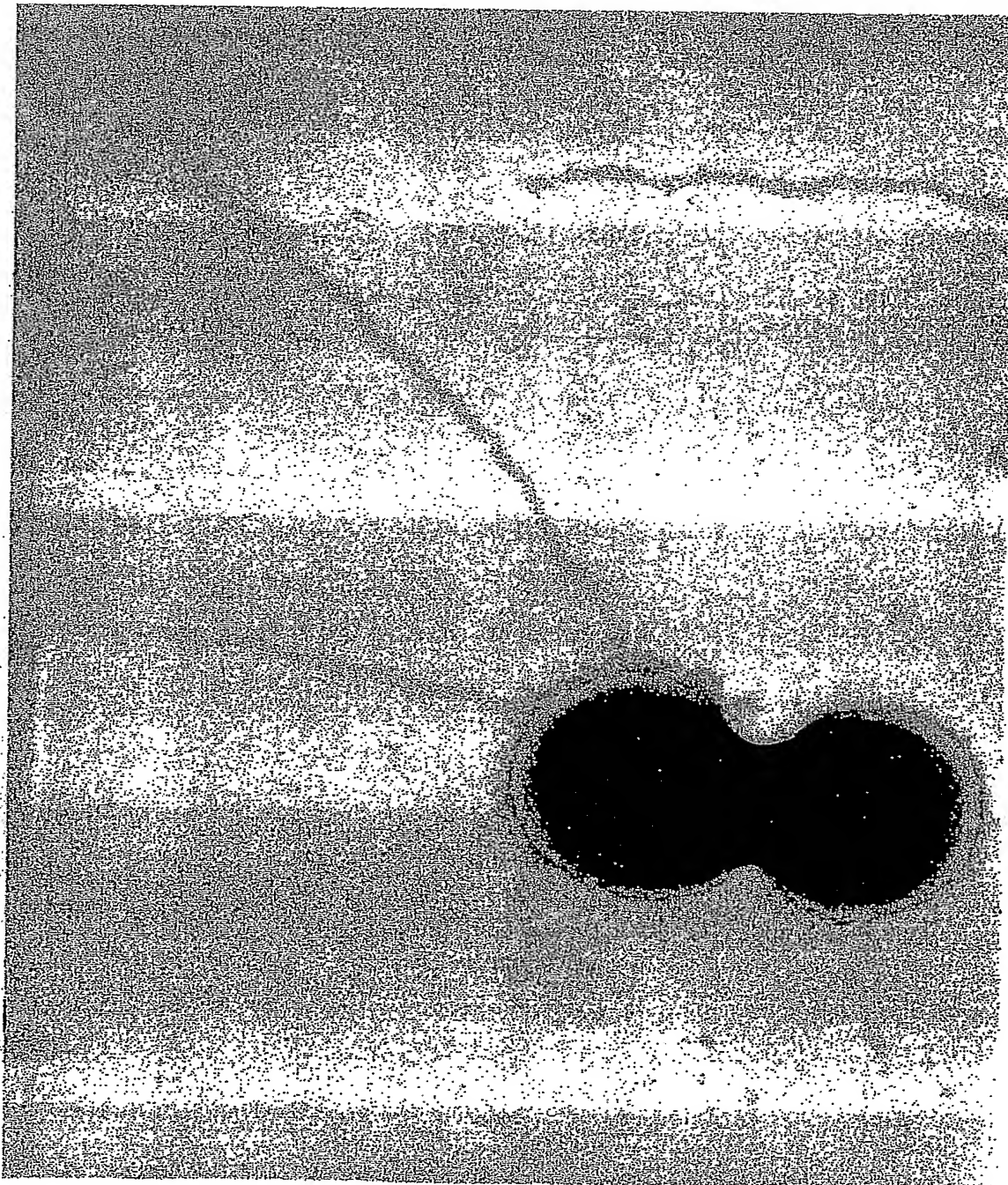
PHD :

Alpha helix	(Hh) :	148 is 16.43%	3 ₁₀ helix	(Gg) :	0 is 0.00%
Pi helix	(Ii) :	0 is 0.00%	Beta bridge	(Bb) :	0 is 0.00%
Extended strand	(Ee) :	243 is 26.97%	Beta turn	(Tt) :	0 is 0.00%
Bend region	(Ss) :	0 is 0.00%	Random coil	(Cc) :	510 is 56.60%
Ambiguous states (?) :	0 is 0.00%		Other states		0 is 0.00%



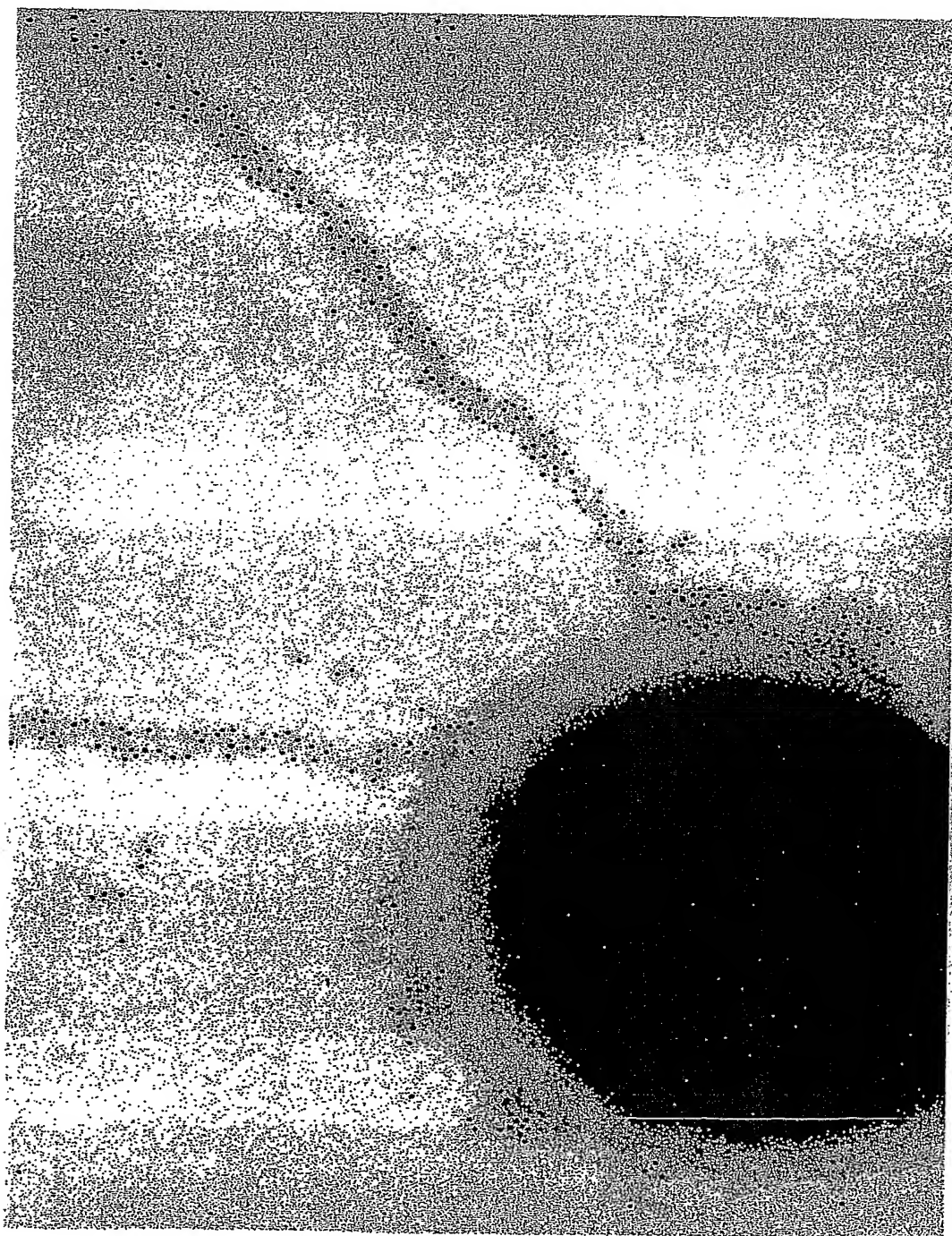
PCT/US05/27239/13/487

Figure 13



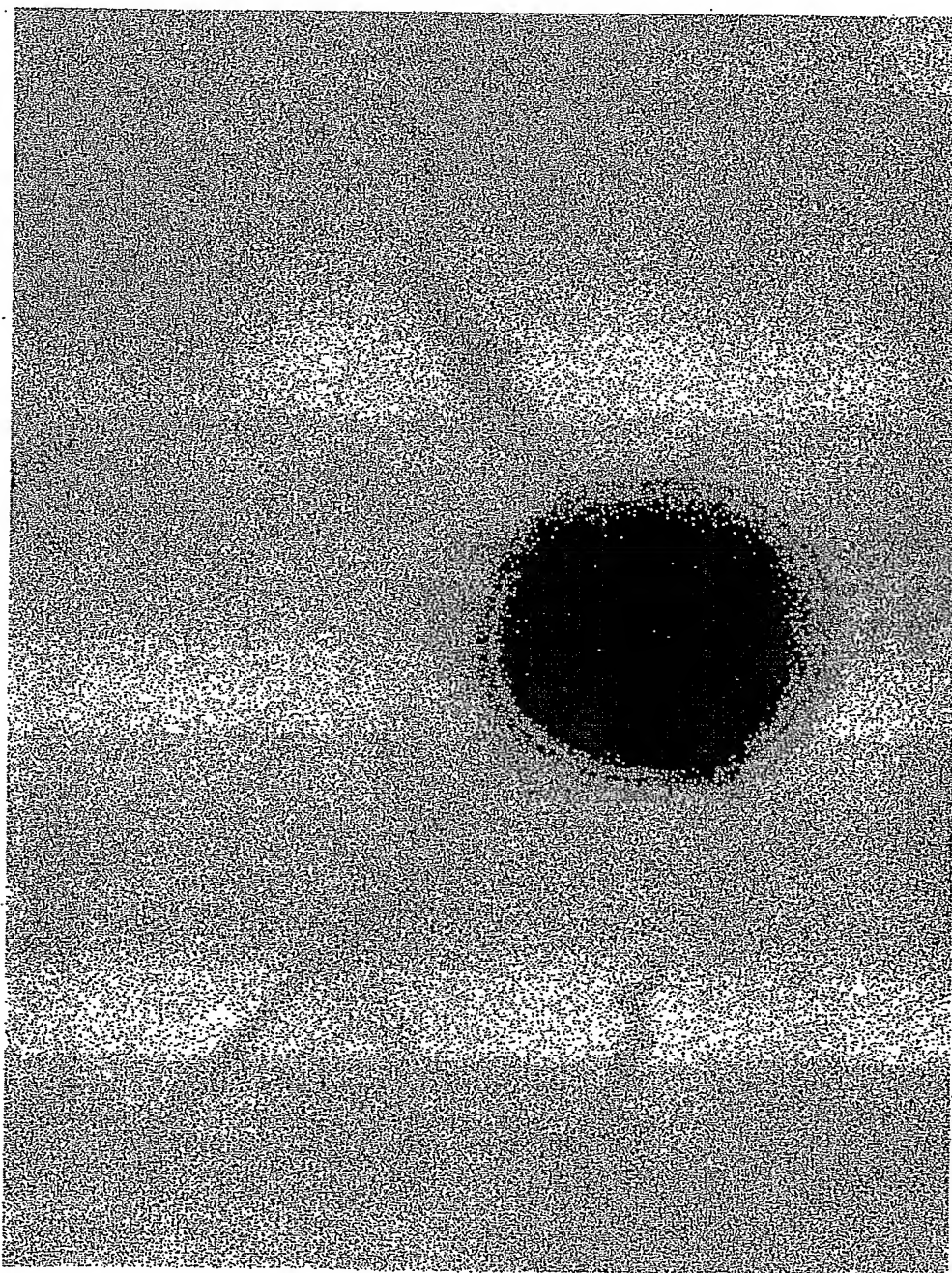
PCT/US05/27239 14/487

Figure 14



PCT/US05/27239 15/487

Figure 15



PCT/US05/27239 16/487

Figure 16

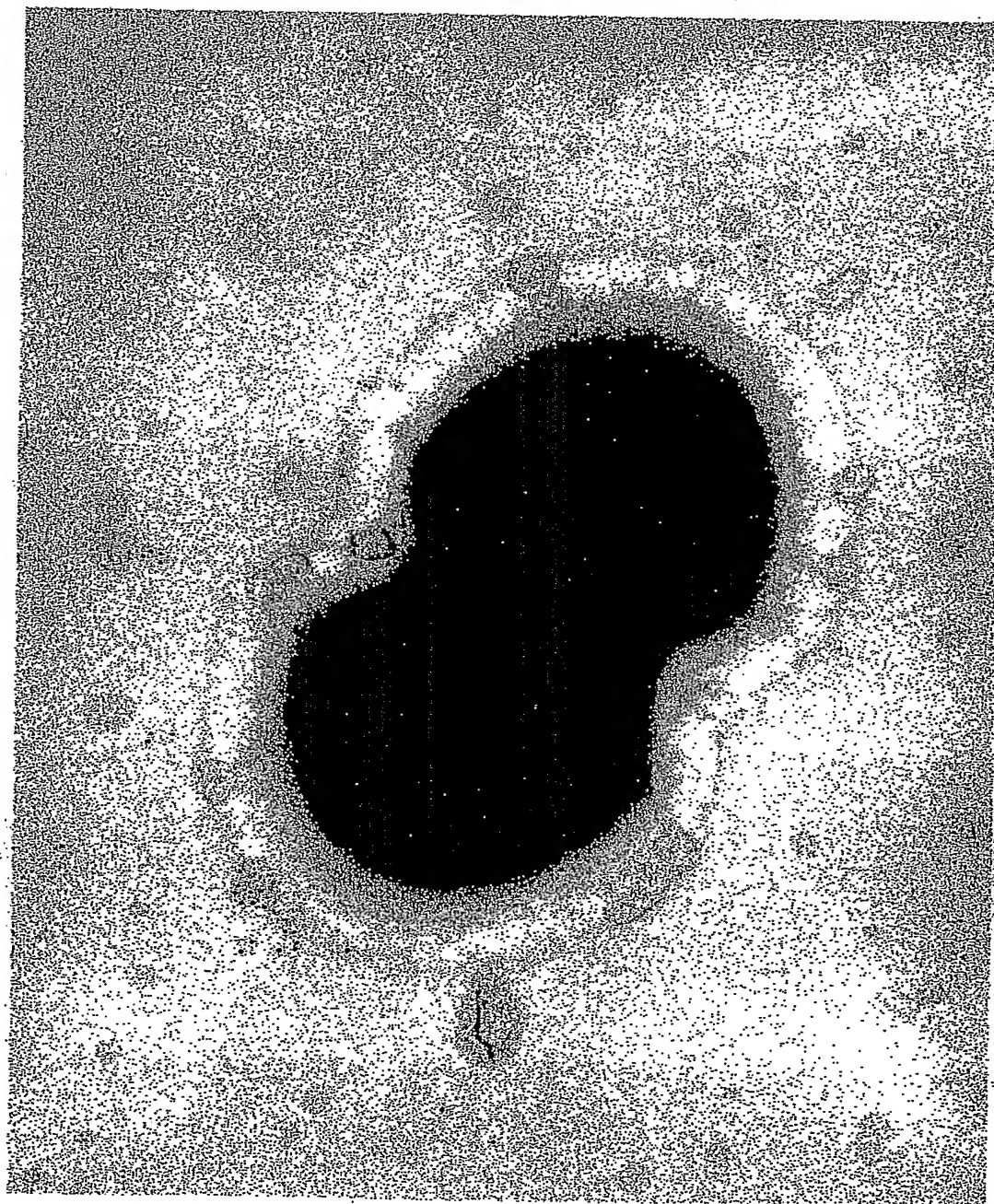
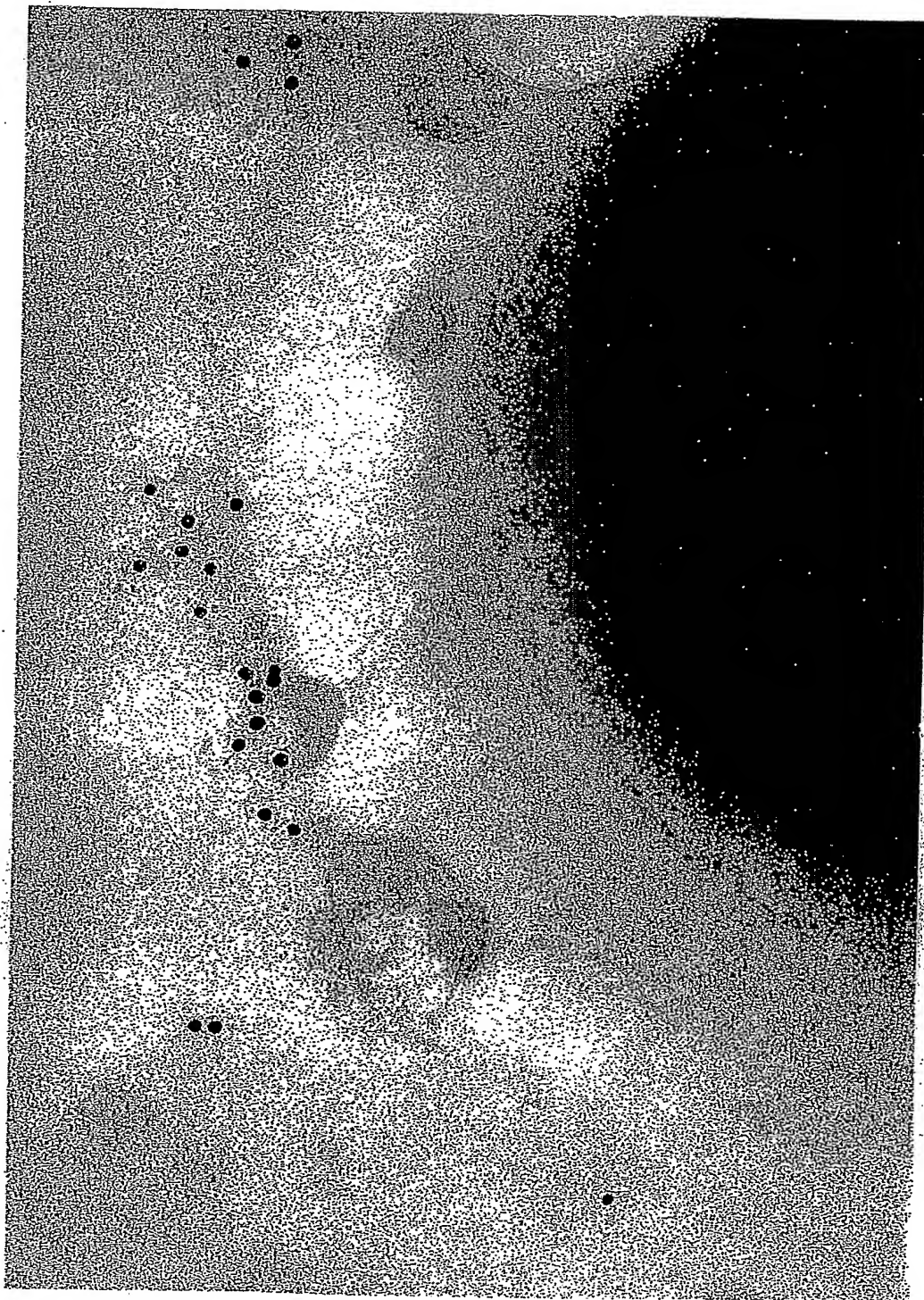
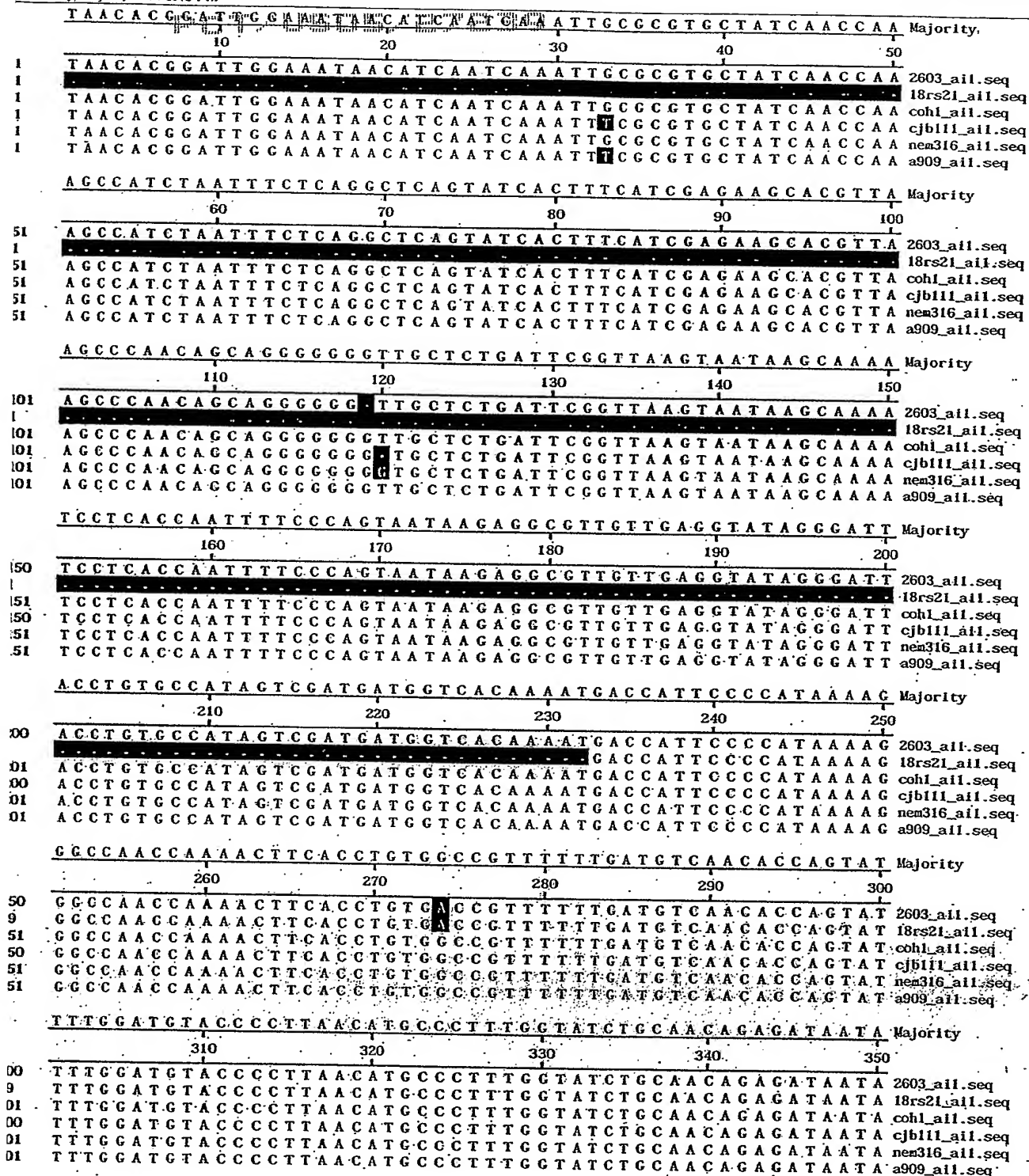


Figure 17



18/487



19/487

T G A C C C A A A G A A G A A T C A C C G A T G A C T T T A A C C G T A A T C T T G C T A T C G C C Majority						
	360	370	380	390	400	
150	T G A C C A A A G G A A G A A T C A C C G A T G A C T T T A A C C G T A A T C T T G C T A T C G C C					2603_all.seq
119	T G A C C A A A G G A A G A A T C A C C G A T G A C T T T A A C C G T A A T C T T G C T A T C G C C					18rs21_all.seq
151	T G A C C A A A G G A A G A A T C A C C G A T G A C T T T A A C C G T A A T C T T G C T A T C G C C					cohl_all.seq
150	T G A C C A A A G G A A G A A T C A C C G A T G A C T T T A A C C G T A A T C T T G C T A T C G C C					cjb111_all.seq
151	T G A C C A A A G G A A G A A T C A C C G A T G A C T T T A A C C G T A A T C T T G C T A T C G C C					nem316_all.seq
151	T G A C C A A A G G A A G A A T C A C C G A T G A C T T T A A C C G T A A T C T T G C T A T C G C C					a909_all.seq
T T T T T G A T T A G C C G C T A A T A T T T T G A T T A G C A A T C A G G G T G C C A C C A A G A G Majority						
	410	420	430	440	450	
100	T T T T T G A T T A G C C G C T A A T A T T T T G A T T A G C A A T C A G G G T G C C A C C A A G A G					2603_all.seq
69	T T T T T G A T T A G C C G C T A A T A T T T T G A T T A G C A A T C A G G G T G C C A C C A A G A G					18rs21_all.seq
101	T T T T T G A T T A G C C G C T A A T A T T T T G A T T A G C A A T C A G G G T G C C A C C A A G A G					cohl_all.seq
100	T T T T T G A T T A G C C G C T A A T A T T T T G A T T A G C A A T C A G G G T G C C A C C A A G A G					cjb111_all.seq
101	T T T T T G A T T A G C C G C T A A T A T T T T G A T T A G C A A T C A G G G T G C C A C C A A G A G					nem316_all.seq
101	T T T T T G A T T A G C C G C T A A T A T T T T G A T T A G C A A T C A G G G T G C C A C C A A G A G					a909_all.seq
C A A C T G T T G A A G A C G A C A A G G T A T G A T G T T T T T C C T G A G C C A A T T T A A C A Majority						
	460	470	480	490	500	
50	C A A C T G T T G A A G A C G A C A A G G T A T G A T G T T T T T C C T G A G C C A A T T T A A C A					2603_all.seq
19	C A A C T G T T G A A G A C G A C A A G G T A T G A T G T T T T T C C T G A G C C A A T T T A A C A					18rs21_all.seq
51	C A A C T G T T G A A G A C G A C A A G G T A T G A T G T T T T T C C T G A G C C A A T T T A A C A					cohl_all.seq
50	C A A C T G T T G A A G A C G A C A A G G T A T G A T G T T T T T C C T G A G C C A A T T T A A C A					cjb111_all.seq
51	C A A C T G T T G A A G A C G A C A A G G T A T G A T G T T T T T C C T G A G C C A A T T T A A C A					nem316_all.seq
51	C A A C T G T T G A A G A C G A C A A G G T A T G A T G T T T T T C C T G A G C C A A T T T A A C A					a909_all.seq
G T C T C G G T G C T A T C G A G G A C A T A A G A C C T G A A G G C A C C A G A G G C A G A A A T Majority						
	510	520	530	540	550	
00	G T C T C G G T G C T A T C G A G G A C A T A A G A C C T G A A G G C A C C A G A G G C A G A A A T					2603_all.seq
69	G T C T C G G T G C T A T C G A G G A C A T A A G A C C T G A A G G C A C C A G A G G C A G A A A T					18rs21_all.seq
01	G T C T C G G T G C T A T C G A G G A C A T A A G A C C T G A A G G C A C C A G A G G C A G A A A T					cohl_all.seq
00	G T C T C G G T G C T A T C G A G G A C A T A A G A C C T G A A G G C A C C A G A G G C A G A A A T					cjb111_all.seq
01	G T C T C G G T G C T A T C G A G G A C A T A A G A C C T G A A G G C A C C A G A G G C A G A A A T					nem316_all.seq
01	G T C T C G G T G C T A T C G A G G A C A T A A G A C C T G A A G G C A C C A G A G G C A G A A A T					a909_all.seq
T G A T T T A A T A A T T T T A T C C A T A A C C T A T G T T A T A G C A C A A A G A G A G A G T T Majority						
	560	570	580	590	600	
50	T G A T T T A A T A A T T T T A T C C A T A A C C T A T G T T A T A G C A C A A A G A G A G A G T T					2603_all.seq
19	T G A T T T A A T A A T T T T A T C C A T A A C C T A T G T T A T A G C A C A A A G A G A G A G T T					18rs21_all.seq
51	T G A T T T A A T A A T T T T A T C C A T A A C C T A T G T T A T A G C A C A A A G A G A G A G T T					cohl_all.seq
50	T G A T T T A A T A A T T T T A T C C A T A A C C T A T G T T A T A G C A C A A A G A G A G A G T T					cjb111_all.seq
51	T G A T T T A A T A A T T T T A T C C A T A A C C T A T G T T A T A G C A C A A A G A G A G A G T T					nem316_all.seq
51	T G A T T T A A T A A T T T T A T C C A T A A C C T A T G T T A T A G C A C A A A G A G A G A G T T					a909_all.seq
T G T A T G G A C T G G A T T A G C T G A A A A T T T T A G A C T A A A A A G T A G T G C C A G T G Majority						
	610	620	630	640	650	
10	T G T A T G G A C T G G A T T A G C T G A A A A T T T T A G A C T A A A A A G T A G T G C C A G T G					2603_all.seq
19	T G T A T G G A C T G G A T T A G C T G A A A A T T T T A G A C T A A A A A G T A G T G C C A G T G					18rs21_all.seq
11	T G T A T G G A C T G G A T T A G C T G A A A A T T T T A G A C T A A A A A G T A G T G C C A G T G					cohl_all.seq
10	T G T A T G G A C T G G A T T A G C T G A A A A T T T T A G A C T A A A A A G T A G T G C C A G T G					cjb111_all.seq
11	T G T A T G G A C T G G A T T A G C T G A A A A T T T T A G A C T A A A A A G T A G T G C C A G T G					nem316_all.seq
11	T G T A T G G A C T G G A T T A G C T G A A A A T T T T A G A C T A A A A A G T A G T G C C A G T G					a909_all.seq
G A T T T T G A A A G G T A A C A A A A G A G T C T T C C G T T T A A A C G A G A A G C C T G T T T Majority						
	660	670	680	690	700	
10	G A T T T T G A A A G G T A A C A A A A G A G T C T T C C G T T T A A A C G A G A A G C C T G T T T					2603_all.seq
19	G A T T T T G A A A G G T A A C A A A A G A G T C T T C C G T T T A A A C G A G A A G C C T G T T T					18rs21_all.seq
11	G A T T T T G A A A G G T A A C A A A A G A G T C T T C C G T T T A A A C G A G A A G C C T G T T T					cohl_all.seq
10	G A T T T T G A A A G G T A A C A A A A G A G T C T T C C G T T T A A A C G A G A A G C C T G T T T					cjb111_all.seq
11	G A T T T T G A A A G G T A A C A A A A G A G T C T T C C G T T T A A A C G A G A A G C C T G T T T					nem316_all.seq
11	G A T T T T G A A A G G T A A C A A A A G A G T C T T C C G T T T A A A C G A G A A G C C T G T T T					a909_all.seq

FIGURE 18 A

TTAGTCGTTTTCAGATGAAGGCCAAAAGGACGATGAATAATTCCGTACCTT Majority									
710		720		730		740		750	
00	TTAGTCGTTTTCAGATGAAGGCCAAAAGGACGATGAATAATTCCGTACCTT	2603_all.seq							
69	TTAGTCGTTTTCAGATGAAGGCCAAAAGGACGATGAATAATTCCGTACCTT	18rs21_all.seq							
01	TTAGTCGTTTTCAGATGAAGGCCAAAAGGACGATGAATAATTCCGTACCTT	cohl_all.seq							
00	TTAGTCGTTTTCAGATGAAGGCCAAAAGGACGATGAATAATTCCGTACCTT	cjb111_all.seq							
01	TTAGTCGTTTTCAGATGAAGGCCAAAAGGACGATGAATAATTCCGTACCTT	nem316_all.seq							
01	TTAGTCGTTTTCAGATGAAGGCCAAAAGGACGATGAATAATTCCGTACCTT	a909_all.seq							
CATGGATTGCTATGTTACTGGCATGAGGTCTCACCAGATATTTTAGTAAGATA Majority									
760		770		780		790		800	
50	CATGGATTGCTATGTTACTGGCATGAGGTCTCACCAGATATTTTAGTAAGATA	2603_all.seq							
19	CATGGATTGCTATGTTACTGGCATGAGGTCTCACCAGATATTTTAGTAAGATA	18rs21_all.seq							
51	CATGGATTGCTATGTTACTGGCATGAGGTCTCACCAGATATTTTAGTAAGATA	cohl_all.seq							
50	CATGGATTGCTATGTTACTGGCATGAGGTCTCACCAGATATTTTAGTAAGATA	cjb111_all.seq							
51	CATGGATTGCTATGTTACTGGCATGAGGTCTCACCAGATATTTTAGTAAGATA	nem316_all.seq							
51	CATGGATTGCTATGTTACTGGCATGAGGTCTCACCAGATATTTTAGTAAGATA	a909_all.seq							
TTCGTTTGAAGATATTTCCACGTAATTTTTTAAAGGTTTTTAAGAAAATCTG Majority									
810		820		830		840		850	
00	TTCGTTTGAAGATATTTCCACGTAATTTTTTAAAGGTTTTTAAGAAAATCTG	2603_all.seq							
69	TTCGTTTGAAGATATTTCCACGTAATTTTTTAAAGGTTTTTAAGAAAATCTG	18rs21_all.seq							
01	TTCGTTTGAAGATATTTCCACGTAATTTTTTAAAGGTTTTTAAGAAAATCTG	cohl_all.seq							
00	TTCGTTTGAAGATATTTCCACGTAATTTTTTAAAGGTTTTTAAGAAAATCTG	cjb111_all.seq							
01	TTCGTTTGAAGATATTTCCACGTAATTTTTTAAAGGTTTTTAAGAAAATCTG	nem316_all.seq							
01	TTCGTTTGAAGATATTTCCACGTAATTTTTTAAAGGTTTTTAAGAAAATCTG	a909_all.seq							
TGGTGTCGTAAAAAATGTAATAAATTTCCGCTACTTCCCAATCCGTACCCCT Majority									
860		870		880		890		900	
50	TGGTGTCGTAAAAAATGTAATAAATTTCCGCTACTTCCCAATCCGTACCCCT	2603_all.seq							
19	TGGTGTCGTAAAAAATGTAATAAATTTCCGCTACTTCCCAATCCGTACCCCT	18rs21_all.seq							
51	TGGTGTCGTAAAAAATGTAATAAATTTCCGCTACTTCCCAATCCGTACCCCT	cohl_all.seq							
50	TGGTGTCGTAAAAAATGTAATAAATTTCCGCTACTTCCCAATCCGTACCCCT	cjb111_all.seq							
51	TGGTGTCGTAAAAAATGTAATAAATTTCCGCTACTTCCCAATCCGTACCCCT	nem316_all.seq							
51	TGGTGTCGTAAAAAATGTAATAAATTTCCGCTACTTCCCAATCCGTACCCCT	a909_all.seq							
CTTTGCAGTAGTAATTGTCCTCCTTAATTTTTGCTTTTAGAATATAAATCT Majority									
910		920		930		940		950	
00	CTTTGCAGTAGTAATTGTCCTCCTTAATTTTTGCTTTTAGAATATAAATCT	2603_all.seq							
19	CTTTGCAGTAGTAATTGTCCTCCTTAATTTTTGCTTTTAGAATATAAATCT	18rs21_all.seq							
01	CTTTGCAGTAGTAATTGTCCTCCTTAATTTTTGCTTTTAGAATATAAATCT	cohl_all.seq							
00	CTTTGCAGTAGTAATTGTCCTCCTTAATTTTTGCTTTTAGAATATAAATCT	cjb111_all.seq							
01	CTTTGCAGTAGTAATTGTCCTCCTTAATTTTTGCTTTTAGAATATAAATCT	nem316_all.seq							
01	CTTTGCAGTAGTAATTGTCCTCCTTAATTTTTGCTTTTAGAATATAAATCT	a909_all.seq							
TTGCAAGCAAATGTCAGAGTATTTTTTTAAAAACTGAGCGTAAGTCCGAAT Majority									
960		970		980		990		1000	
00	TTGCAAGCAAATGTCAGAGTATTTTTTTAAAAACTGAGCGTAAGTCCGAAT	2603_all.seq							
19	TTGCAAGCAAATGTCAGAGTATTTTTTTAAAAACTGAGCGTAAGTCCGAAT	18rs21_all.seq							
01	TTGCAAGCAAATGTCAGAGTATTTTTTTAAAAACTGAGCGTAAGTCCGAAT	cohl_all.seq							
00	TTGCAAGCAAATGTCAGAGTATTTTTTTAAAAACTGAGCGTAAGTCCGAAT	cjb111_all.seq							
01	TTGCAAGCAAATGTCAGAGTATTTTTTTAAAAACTGAGCGTAAGTCCGAAT	nem316_all.seq							
01	TTGCAAGCAAATGTCAGAGTATTTTTTTAAAAACTGAGCGTAAGTCCGAAT	a909_all.seq							
CTGAGATATATAGGTAAGTTGGCAATATCAGATACTTTGAGTTTGGAGTAG Majority									
1010		1020		1030		1040		1050	
00	CTGAGATATATAGGTAAGTTGGCAATATCAGATACTTTGAGTTTGGAGTAG	2603_all.seq							
19	CTGAGATATATAGGTAAGTTGGCAATATCAGATACTTTGAGTTTGGAGTAG	18rs21_all.seq							
01	CTGAGATATATAGGTAAGTTGGCAATATCAGATACTTTGAGTTTGGAGTAG	cohl_all.seq							
00	CTGAGATATATAGGTAAGTTGGCAATATCAGATACTTTGAGTTTGGAGTAG	cjb111_all.seq							
01	CTGAGATATATAGGTAAGTTGGCAATATCAGATACTTTGAGTTTGGAGTAG	nem316_all.seq							
01	CTGAGATATATAGGTAAGTTGGCAATATCAGATACTTTGAGTTTGGAGTAG	a909_all.seq							

FIGURE 18 B

Alignment Report of A1-1 alignment, using J. Hein method with Weighted residue weight table.
Thursday, July 29, 2004 5:46 PM

A G G T C G G T T G T C C A C A T A A T G G A C A A T A C T A T T G T A C A T T T G C T G C T T G T C Majority									
1060		1070		1080		1090		1100	
1050	A	G	G	T	C	G	G	T	T
819	A	G	G	T	C	G	G	T	T
1051	A	G	G	T	C	G	G	T	T
1050	A	G	G	T	C	G	G	T	T
1051	A	G	G	T	C	G	G	T	T
1051	A	G	G	T	C	G	G	T	T
A G A G A T G C T C T T A T T G C T T A A G G A T T C T G A A A A A T C A A T A A G A G C T G C A C Majority									
1110		1120		1130		1140		1150	
1100	A	G	A	G	A	T	G	C	T
869	A	G	A	G	A	T	G	C	T
1101	A	G	A	G	A	T	G	C	T
1100	A	G	A	G	A	T	G	C	T
1101	A	G	A	G	A	T	G	C	T
1101	A	G	A	G	A	T	G	C	T
A G C G A A T T C T T G A A A C A T C A A T A A G A T C A G G A G C C T C T T C G T T T A A A G C C Majority									
1160		1170		1180		1190		1200	
1150	A	G	C	G	A	A	T	T	C
919	A	G	C	G	A	A	T	T	C
1151	A	G	C	G	A	A	T	T	C
1150	A	G	C	G	A	A	T	T	C
1151	A	G	C	G	A	A	T	T	C
1151	A	G	C	G	A	A	T	T	C
A T A T A G T G C T T T A C C A G C G C A T A A C T T T T A G C C A C A T C A G T A T T T T C C T C Majority									
1210		1220		1230		1240		1250	
1200	A	T	A	T	A	G	T	G	C
169	A	T	A	T	A	G	T	G	C
1201	A	T	A	T	A	G	T	G	C
1200	A	T	A	T	A	G	T	G	C
1201	A	T	A	T	A	G	T	G	C
1201	A	T	A	T	A	G	T	G	C
G A A A C T T A A T T C T A G T A A T T T T G T T A A G T A A A C A A C A G T T A A G T T C T T T T Majority									
1260		1270		1280		1290		1300	
250	G	A	A	A	C	T	T	A	A
019	G	A	A	A	C	T	T	A	A
251	G	A	A	A	C	T	T	A	A
250	G	A	A	A	C	T	T	A	A
251	G	A	A	A	C	T	T	A	A
251	G	A	A	A	C	T	T	A	A
C A G C T C T T A G G G C A G G G A T T G A A G A T G A G C T A A C A C T G G A T G A T G G G A G G Majority									
1310		1320		1330		1340		1350	
300	C	A	G	C	T	C	T	T	A
069	C	A	G	C	T	C	T	T	A
301	C	A	G	C	T	C	T	T	A
300	C	A	G	C	T	C	T	T	A
301	C	A	G	C	T	C	T	T	A
301	C	A	G	C	T	C	T	T	A
C G A T T A A T T T C T T G C T T T A A C A G T T G A G T G T T A C C C A G C T T A A C G A G A T C Majority									
1360		1370		1380		1390		1400	
350	C	G	A	T	T	A	A	T	T
119	C	G	A	T	T	A	A	T	T
351	C	G	A	T	T	A	A	T	T
350	C	G	A	T	T	A	A	T	T
351	C	G	A	T	T	A	A	T	T
351	C	G	A	T	T	A	A	T	T

FIGURE 18 C

22/487

WO 2006/078318

Alignment Report of Al-1 alignment, using J. Hein method with Weighted residue weight table.
Thursday, July 29, 2004 5:46 PM

PCT/US2005/027239

Page 5

A A T A A T G T C A T T C A G A T G G T T A A A A C A G T C G G T A A C T G A A A A G A G T T T T Majority																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
1410										1420										1430										1440										1450																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
400	A	A	T	A	A	T	G	T	C	A	T	T	C	A	G	A	T	G	G	T	T	A	A	A	C	A	G	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G	G	T	A	A	C	T	G

23/487

G A C A G A C A G G A A A A T T A A G A A T G A T T C T T C A A A A A G A C C C T C A T A A A C A Majority										
1760		1770		1780		1790		1800		
750	G A C A G A C A G G A A A A T T A A G A A T G A T T C T T C A A A A A G A C C C T C A T A A A C A									2603_all.seq
519	G A C A G A C A G G A A A A T T A A G A A T G A T T C T T C A A A A A G A C C C T C A T A A A C A									18rs21_all.seq
751	G A C A G A C A G G A A A A T T A A G A A T G A T T C T T C A A A A A G A C C C T C A T A A A C A									coh1_all.seq
750	G A C A G A C A G G A A A A T T A A G A A T G A T T C T T C A A A A A G A C C C T C A T A A A C A									cjb111_all.seq
751	G A C A G A C A G G A A A A T T A A G A A T G A T T C T T C A A A A A G A C C C T C A T A A A C A									nem316_all.seq
751	G A C A G A C A G G A A A A T T A A G A A T G A T T C T T C A A A A A G A C C C T C A T A A A C A									a909_all.seq
G T G A T A T C T T G G T T A T A A G G G A T A G C T A A A T G T T T T A A A A A C T G A T A C T A Majority										
1810		1820		1830		1840		1850		
800	G T G A T A T C T T G G T T A T A A G G G A T A G C T A A A T G T T T T A A A A A C T G A T A C T A									2603_all.seq
569	G T G A T A T C T T G G T T A T A A G G G A T A G C T A A A T G T T T T A A A A A C T G A T A C T A									18rs21_all.seq
801	G T G A T A T C T T G G T T A T A A G G G A T A G C T A A A T G T T T T A A A A A C T G A T A C T A									coh1_all.seq
800	G T G A T A T C T T G G T T A T A A G G G A T A G C T A A A T G T T T T A A A A A C T G A T A C T A									cjb111_all.seq
801	G T G A T A T C T T G G T T A T A A G G G A T A G C T A A A T G T T T T A A A A A C T G A T A C T A									nem316_all.seq
801	G T G A T A T C T T G G T T A T A A G G G A T A G C T A A A T G T T T T A A A A A C T G A T A C T A									a909_all.seq
A G G C A A C A G A T A G T C T T C G T T A C C A T A T A A C T G A A C G A G T T C C T T G T C T C Majority										
1860		1870		1880		1890		1900		
850	A G G C A A C A G A T A G T C T T C G T T A C C A T A T A A C T G A A C G A G T T C C T T G T C T C									2603_all.seq
619	A G G C A A C A G A T A G T C T T C G T T A C C A T A T A A C T G A A C G A G T T C C T T G T C T C									18rs21_all.seq
851	A G G C A A C A G A T A G T C T T C G T T A C C A T A T A A C T G A A C G A G T T C C T T G T C T C									coh1_all.seq
850	A G G C A A C A G A T A G T C T T C G T T A C C A T A T A A C T G A A C G A G T T C C T T G T C T C									cjb111_all.seq
851	A G G C A A C A G A T A G T C T T C G T T A C C A T A T A A C T G A A C G A G T T C C T T G T C T C									nem316_all.seq
851	A G G C A A C A G A T A G T C T T C G T T A C C A T A T A A C T G A A C G A G T T C C T T G T C T C									a909_all.seq
G T G A C A T C A C T G A A A T A G G T A G T T G A G A T A T G G T A T G C A A T G T T T G A A C A Majority										
1910		1920		1930		1940		1950		
900	G T G A C A T C A C T G A A A T A G G T A G T T G A G A T A T G G T A T G C A A T G T T T G A A C A									2603_all.seq
669	G T G A C A T C A C T G A A A T A G G T A G T T G A G A T A T G G T A T G C A A T G T T T G A A C A									18rs21_all.seq
901	G T G A C A T C A C T G A A A T A G G T A G T T G A G A T A T G G T A T G C A A T G T T T G A A C A									coh1_all.seq
900	G T G A C A T C A C T G A A A T A G G T A G T T G A G A T A T G G T A T G C A A T G T T T G A A C A									cjb111_all.seq
901	G T G A C A T C A C T G A A A T A G G T A G T T G A G A T A T G G T A T G C A A T G T T T G A A C A									nem316_all.seq
901	G T G A C A T C A C T G A A A T A G G T A G T T G A G A T A T G G T A T G C A A T G T T T G A A C A									a909_all.seq
T G T T T A A A A T C G A A T G T A A C C A T T T G A T A G A C C G C C T T C A T T A T C A T T T C Majority										
1960		1970		1980		1990		2000		
950	T G T T T A A A A T C G A A T G T A A C C A T T T G A T A G A C C G C C T T C A T T A T C A T T T C									2603_all.seq
719	T G T T T A A A A T C G A A T G T A A C C A T T T G A T A G A C C G C C T T C A T T A T C A T T T C									18rs21_all.seq
951	T G T T T A A A A T C G A A T G T A A C C A T T T G A T A G A C C G C C T T C A T T A T C A T T T C									coh1_all.seq
950	T G T T T A A A A T C G A A T G T A A C C A T T T G A T A G A C C G C C T T C A T T A T C A T T T C									cjb111_all.seq
951	T G T T T A A A A T C G A A T G T A A C C A T T T G A T A G A C C G C C T T C A T T A T C A T T T C									nem316_all.seq
951	T G T T T A A A A T C G A A T G T A A C C A T T T G A T A G A C C G C C T T C A T T A T C A T T T C									a909_all.seq
T A G A A T T T T T C T T T A G G T T T G T A A A G A C T A C A A A A T A A A A T G A T G A A A A C Majority										
2010		2020		2030		2040		2050		
300	T A G A A T T T T T C T T T A G G T T T G T A A A G A C T A C A A A A T A A A A T G A T G A A A A C									2603_all.seq
769	T A G A A T T T T T C T T T A G G T T T G T A A A G A C T A C A A A A T A A A A T G A T G A A A A C									18rs21_all.seq
301	T A G A A T T T T T C T T T A G G T T T G T A A A G A C T A C A A A A T A A A A T G A T G A A A A C									coh1_all.seq
300	T A G A A T T T T T C T T T A G G T T T G T A A A G A C T A C A A A A T A A A A T G A T G A A A A C									cjb111_all.seq
301	T A G A A T T T T T C T T T A G G T T T G T A A A G A C T A C A A A A T A A A A T G A T G A A A A C									nem316_all.seq
301	T A G A A T T T T T C T T T A G G T T T G T A A A G A C T A C A A A A T A A A A T G A T G A A A A C									a909_all.seq
A A C T A T C T T G T G G A T A C A C T A A A A A G A C A C G C T A A T T A G C A A A C T C T C T C Majority										
2060		2070		2080		2090		2100		
350	A A C T A T C T T G T G G A T A C A C T A A A A A G A C A C G C T A A T T A G C A A A C T C T C T C									2603_all.seq
319	A A C T A T C T T G T G G A T A C A C T A A A A A G A C A C G C T A A T T A G C A A A C T C T C T C									18rs21_all.seq
351	A A C T A T C T T G T G G A T A C A C T A A A A A G A C A C G C T A A T T A G C A A A C T C T C T C									coh1_all.seq
350	A A C T A T C T T G T G G A T A C A C T A A A A A G A C A C G C T A A T T A G C A A A C T C T C T C									cjb111_all.seq
351	A A C T A T C T T G T G G A T A C A C T A A A A A G A C A C G C T A A T T A G C A A A C T C T C T C									nem316_all.seq
351	A A C T A T C T T G T G G A T A C A C T A A A A A G A C A C G C T A A T T A G C A A A C T C T C T C									a909_all.seq

FIGURE 18 E

TTCATCATCTCTCTCACCATTATTATACTACTATTTTATATGACAAATAAAGG										Majority
2110		2120		2130		2140		2150		
2100	TTCATCATCTCTCTCACCATTATTATACTACTATTTTATATGACAAATAAAGG	2603_all.seq								
1869	TTCATCATCTCTCTCACCATTATTATACTACTATTTTATATGACAAATAAAGG	18rs21_all.seq								
2101	TTCATCATCTCTCTCACCATTATTATACTACTATTTTATATGACAAATAAAGG	cohl1_all.seq								
2100	TTCATCATCTCTCTCACCATTATTATACTACTATTTTATATGACAAATAAAGG	cjb111_all.seq								
2101	TTCATCATCTCTCTCACCATTATTATACTACTATTTTATATGACAAATAAAGG	nem316_all.seq								
2101	TTCATCATCTCTCTCACCATTATTATACTACTATTTTATATGACAAATAAAGG	a909_all.seq								
TGATTTTGTGTTAAAAATATAA CTTTGAAAATCCACATATATTTTAAATCTT										Majority
2160		2170		2180		2190		2200		
2150	TGATTTTGTGTTAAAAATATAA CTTTGAAAATCCACATATATTTTAAATCTT	2603_all.seq								
1919	TGATTTTGTGTTAAAAATATAA CTTTGAAAATCCACATATATTTTAAATCTT	18rs21_all.seq								
2151	TGATTTTGTGTTAAAAATATAA CTTTGAAAATCCACATATATTTTAAATCTT	cohl1_all.seq								
2150	TGATTTTGTGTTAAAAATATAA CTTTGAAAATCCACATATATTTTAAATCTT	cjb111_all.seq								
2151	TGATTTTGTGTTAAAAATATAA CTTTGAAAATCCACATATATTTTAAATCTT	nem316_all.seq								
2151	TGATTTTGTGTTAAAAATATAA CTTTGAAAATCCACATATATTTTAAATCTT	a909_all.seq								
CCGCTCTGAAAAAA - TAAATAAAAAATAGTAAAAAATAAACACGAATTTAAAAA										Majority
2210		2220		2230		2240		2250		
2199	CCGCTCTGAAAAAA - TAAATAAAAAATAGTAAAAAATAAACACGAATTTAAAAA	2603_all.seq								
1968	CCGCTCTGAAAAAA - TAAATAAAAAATAGTAAAAAATAAACACGAATTTAAAAA	18rs21_all.seq								
2200	CCGCTCTGAAAAAA - TAAATAAAAAATAGTAAAAAATAAACACGAATTTAAAAA	cohl1_all.seq								
2200	CCGCTCTGAAAAAA - TAAATAAAAAATAGTAAAAAATAAACACGAATTTAAAAA	cjb111_all.seq								
2201	CCGCTCTGAAAAAA - TAAATAAAAAATAGTAAAAAATAAACACGAATTTAAAAA	nem316_all.seq								
2201	CCGCTCTGAAAAAA - TAAATAAAAAATAGTAAAAAATAAACACGAATTTAAAAA	a909_all.seq								
TAAGCCAAATTTTAAAGAAAAATCTGTGCTAAACTTTAATAGTTTTGTGCT										Majority
2260		2270		2280		2290		2300		
248	TAAGCCAAATTTTAAAGAAAAATCTGTGCTAAACTTTAATAGTTTTGTGCT	2603_all.seq								
2017	TAAGCCAAATTTTAAAGAAAAATCTGTGCTAAACTTTAATAGTTTTGTGCT	18rs21_all.seq								
249	TAAGCCAAATTTTAAAGAAAAATCTGTGCTAAACTTTAATAGTTTTGTGCT	cohl1_all.seq								
249	TAAGCCAAATTTTAAAGAAAAATCTGTGCTAAACTTTAATAGTTTTGTGCT	cjb111_all.seq								
251	TAAGCCAAATTTTAAAGAAAAATCTGTGCTAAACTTTAATAGTTTTGTGCT	nem316_all.seq								
250	TAAGCCAAATTTTAAAGAAAAATCTGTGCTAAACTTTAATAGTTTTGTGCT	a909_all.seq								
TAATAATAATCAGCACTTACAAAGAAACAAAGGGGAAAGCGAGGAGAGAGAAC										Majority
2310		2320		2330		2340		2350		
298	TAATAATAATCAGCACTTACAAAGAAACAAAGGGGAAAGCGAGGAGAGAGAAC	2603_all.seq								
067	TAATAATAATCAGCACTTACAAAGAAACAAAGGGGAAAGCGAGGAGAGAGAAC	18rs21_all.seq								
299	TAATAATAATCAGCACTTACAAAGAAACAAAGGGGAAAGCGAGGAGAGAGAAC	cohl1_all.seq								
299	TAATAATAATCAGCACTTACAAAGAAACAAAGGGGAAAGCGAGGAGAGAGAAC	cjb111_all.seq								
301	TAATAATAATCAGCACTTACAAAGAAACAAAGGGGAAAGCGAGGAGAGAGAAC	nem316_all.seq								
300	TAATAATAATCAGCACTTACAAAGAAACAAAGGGGAAAGCGAGGAGAGAGAAC	a909_all.seq								
TTTTAATGAAATTATCGAAGAAAGTTATTGTTTTTCGGCTGCTGTTTTTAAACA										Majority
2360		2370		2380		2390		2400		
348	TTTTAATGAAATTATCGAAGAAAGTTATTGTTTTTCGGCTGCTGTTTTTAAACA	2603_all.seq								
117	TTTTAATGAAATTATCGAAGAAAGTTATTGTTTTTCGGCTGCTGTTTTTAAACA	18rs21_all.seq								
349	TTTTAATGAAATTATCGAAGAAAGTTATTGTTTTTCGGCTGCTGTTTTTAAACA	cohl1_all.seq								
349	TTTTAATGAAATTATCGAAGAAAGTTATTGTTTTTCGGCTGCTGTTTTTAAACA	cjb111_all.seq								
351	TTTTAATGAAATTATCGAAGAAAGTTATTGTTTTTCGGCTGCTGTTTTTAAACA	nem316_all.seq								
350	TTTTAATGAAATTATCGAAGAAAGTTATTGTTTTTCGGCTGCTGTTTTTAAACA	a909_all.seq								
ATCGGTGGCGGGCTCAACTCTTTGAACCACTAGCTCAGTTTCCGACTCGCAAT										Majority
2410		2420		2430		2440		2450		
398	ATCGGTGGCGGGCTCAACTCTTTGAACCACTAGCTCAGTTTCCGACTCGCAAT	2603_all.seq								
167	ATCGGTGGCGGGCTCAACTCTTTGAACCACTAGCTCAGTTTCCGACTCGCAAT	18rs21_all.seq								
399	ATCGGTGGCGGGCTCAACTCTTTGAACCACTAGCTCAGTTTCCGACTCGCAAT	cohl1_all.seq								
399	ATCGGTGGCGGGCTCAACTCTTTGAACCACTAGCTCAGTTTCCGACTCGCAAT	cjb111_all.seq								
401	ATCGGTGGCGGGCTCAACTCTTTGAACCACTAGCTCAGTTTCCGACTCGCAAT	nem316_all.seq								
400	ATCGGTGGCGGGCTCAACTCTTTGAACCACTAGCTCAGTTTCCGACTCGCAAT	a909_all.seq								

FIGURE 18 F

25/487

C A G T A T T G T A A C A G C T G C A G A A G T G T C A C A A G A A C G C C C A G C G A A A A C A A Majority											
		2460		2470		2480		2490		2500	
2448		C A G T A T T G T A A C A G C T G C A G A A G T G T C A C A A G A A C G C C C A G C G A A A A C A A 2603_all.seq									
2217		C A G T A T T G T A A G A C C T G C A G A A G T G T C A C A A G A A C G C C C A G C G A A A A C A A 18rs21_all.seq									
2449		C A G T A T T G T A A G A C C T G C A G A A G T G T C A C A A G A A C G C C C A G C G A A A A C A A coh1_all.seq									
2449		C A G T A T T G T A A G A C C T G C A G A A G T G T C A C A A G A A C G C C C A G C G A A A A C A A cjb111_all.seq									
2451		C A G T A T T G T A A G A C C T G C A G A A G T G T C A C A A G A A C G C C C A G C G A A A A C A A nem316_all.seq									
2450		C A G T A T T G T A A G A C C T G C A G A A G T G T C A C A A G A A C G C C C A G C G A A A A C A A a909_all.seq									
C A G T A A A T A T C T A T A A A T T A C A A G C T G A T A G T T A T A A A T C G G A A A T T A C T Majority											
		2510		2520		2530		2540		2550	
2498		C A G T A A A T A T C T A T A A A T T A C A A G C T G A T A G T T A T A A A T C G G A A A T T A C T 2603_all.seq									
2267		C A G T A A A T A T C T A T A A A T T A C A A G C T G A T A G T T A T A A A T C G G A A A T T A C T 18rs21_all.seq									
2499		C A G T A A A T A T C T A T A A A T T A C A A G C T G A T A G T T A T A A A T C G G A A A T T A C T coh1_all.seq									
2499		C A G T A A A T A T C T A T A A A T T A C A A G C T G A T A G T T A T A A A T C G G A A A T T A C T cjb111_all.seq									
2501		C A G T A A A T A T C T A T A A A T T A C A A G C T G A T A G T T A T A A A T C G G A A A T T A C T nem316_all.seq									
2500		C A G T A A A T A T C T A T A A A T T A C A A G C T G A T A G T T A T A A A T C G G A A A T T A C T a909_all.seq									
T C T A A T G C T G C T A T C G A G A A T A A A G A C G C C C A A G T A A T A T C T A A C T A T G C Majority											
		2560		2570		2580		2590		2600	
2548		T C T A A T G C T G C T A T C G A G A A T A A A G A C G C C C A A G T A A T A T C T A A C T A T G C 2603_all.seq									
2317		T C T A A T G C T G C T A T C G A G A A T A A A G A C G C C C A A G T A A T A T C T A A C T A T G C 18rs21_all.seq									
2549		T C T A A T G C T G C T A T C G A G A A T A A A G A C G C C C A A G T A A T A T C T A A C T A T G C coh1_all.seq									
2549		T C T A A T G C T G C T A T C G A G A A T A A A G A C G C C C A A G T A A T A T C T A A C T A T G C cjb111_all.seq									
2551		T C T A A T G C T G C T A T C G A G A A T A A A G A C G C C C A A G T A A T A T C T A A C T A T G C nem316_all.seq									
2550		T C T A A T G C T G C T A T C G A G A A T A A A G A C G C C C A A G T A A T A T C T A A C T A T G C a909_all.seq									
T A A A C T T G C T G A C A A T G T A A A A G C T T T G C A A G G T G T A C A G T T T A A A C G T T Majority											
		2610		2620		2630		2640		2650	
598		T A A A C T T G C T G A C A A T G T A A A A G C T T T G C A A G G T G T A C A G T T T A A A C G T T 2603_all.seq									
367		T A A A C T T G C T G A C A A T G T A A A A G C T T T G C A A G G T G T A C A G T T T A A A C G T T 18rs21_all.seq									
599		T A A A C T T G C T G A C A A T G T A A A A G C T T T G C A A G G T G T A C A G T T T A A A C G T T coh1_all.seq									
599		T A A A C T T G C T G A C A A T G T A A A A G C T T T G C A A G G T G T A C A G T T T A A A C G T T cjb111_all.seq									
601		T A A A C T T G C T G A C A A T G T A A A A G C T T T G C A A G G T G T A C A G T T T A A A C G T T nem316_all.seq									
600		T A A A C T T G C T G A C A A T G T A A A A G C T T T G C A A G G T G T A C A G T T T A A A C G T T a909_all.seq									
A T A A A G T C A A G A C G C A T A T T T C T G T T G A T G A A T T G A A A A A T T G A C A A C A Majority											
		2660		2670		2680		2690		2700	
648		A T A A A G T C A A G A C G C A T A T T T C T G T T G A T G A A T T G A A A A A T T G A C A A C A 2603_all.seq									
417		A T A A A G T C A A G A C G C A T A T T T C T G T T G A T G A A T T G A A A A A T T G A C A A C A 18rs21_all.seq									
649		A T A A A G T C A A G A C G C A T A T T T C T G T T G A T G A A T T G A A A A A T T G A C A A C A coh1_all.seq									
549		A T A A A G T C A A G A C G C A T A T T T C T G T T G A T G A A T T G A A A A A T T G A C A A C A cjb111_all.seq									
351		A T A A A G T C A A G A C G C A T A T T T C T G T T G A T G A A T T G A A A A A T T G A C A A C A nem316_all.seq									
350		A T A A A G T C A A G A C G C A T A T T T C T G T T G A T G A A T T G A A A A A T T G A C A A C A a909_all.seq									
G T T G A A G C A G C A G A T G C A A A A G T T G C A A C G A T T C T T G A A C A A G G T G T C A G Majority											
		2710		2720		2730		2740		2750	
198		G T T G A A G C A G C A G A T G C A A A A G T T G C A A C G A T T C T T G A A C A A G G T G T C A G 2603_all.seq									
167		G T T G A A G C A G C A G A T G C A A A A G T T G C A A C G A T T C T T G A A C A A G G T G T C A G 18rs21_all.seq									
199		G T T G A A G C A G C A G A T G C A A A A G T T G C A A C G A T T C T T G A A C A A G G T G T C A G coh1_all.seq									
199		G T T G A A G C A G C A G A T G C A A A A G T T G C A A C G A T T C T T G A A C A A G G T G T C A G cjb111_all.seq									
101		G T T G A A G C A G C A G A T G C A A A A G T T G C A A C G A T T C T T G A A C A A G G T G T C A G nem316_all.seq									
100		G T T G A A G C A G C A G A T G C A A A A G T T G C A A C G A T T C T T G A A C A A G G T G T C A G a909_all.seq									
T C T A C C T C A A A A A A C T A A T G C T C A A G C T T T G C T C G T C G A T G C T C T C G A T T Majority											
		2760		2770		2780		2790		2800	
48		T C T A C C T C A A A A A A C T A A T G C T C A A G C T T T G C T C G T C G A T G C T C T C G A T T 2603_all.seq									
117		T C T A C C T C A A A A A A C T A A T G C T C A A G C T T T G C T C G T C G A T G C T C T C G A T T 18rs21_all.seq									
49		T C T A C C T C A A A A A A C T A A T G C T C A A G C T T T G C T C G T C G A T G C T C T C G A T T coh1_all.seq									
49		T C T A C C T C A A A A A A C T A A T G C T C A A G C T T T G C T C G T C G A T G C T C T C G A T T cjb111_all.seq									
51		T C T A C C T C A A A A A A C T A A T G C T C A A G C T T T G C T C G T C G A T G C T C T C G A T T nem316_all.seq									
50		T C T A C C T C A A A A A A C T A A T G C T C A A G C T T T G C T C G T C G A T G C T C T C G A T T a909_all.seq									

FIGURE 18 G

FIGURE 18 H

27/487

ATTGCTTCGATAAAACAGCTCATAATAGAGATGAGCACTACACTATTGATGAACC Majority									
3160		3170		3180		3190		3200	
148	ATTGCTTCGATAAAACACTGAATAGAGATGAGCACTACACTATTGATGAACC	2603_all.seq							
917	ATTGCTTCGATAAAACACTGAATAGAGATGAGCACTACACTATTGATGAACC	18rs21_all.seq							
149	ATTGCTTCGATAAAACACTGAATAGAGATGAGCACTACACTATTGATGAACC	cohl1_all.seq							
149	ATTGCTTCGATAAAACACTGAATAGAGATGAGCACTACACTATTGATGAACC	cjb111_all.seq							
151	ATTGCTTCGATAAAACACTGAATAGAGATGAGCACTACACTATTGATGAACC	nem316_all.seq							
150	ATTGCTTCGATAAAACACTGAATAGAGATGAGCACTACACTATTGATGAACC	a909_all.seq							
AACAGTTGATAAACCAAAATACATTAAAAATTACGTTTAAACCAGAGAAAT Majority									
3210		3220		3230		3240		3250	
198	AACAGTTGATAAACCAAAATACATTAAAAATTACGTTTAAACCAGAGAAAT	2603_all.seq							
967	AACAGTTGATAAACCAAAATACATTAAAAATTACGTTTAAACCAGAGAAAT	18rs21_all.seq							
199	AACAGTTGATAAACCAAAATACATTAAAAATTACGTTTAAACCAGAGAAAT	cohl1_all.seq							
199	AACAGTTGATAAACCAAAATACATTAAAAATTACGTTTAAACCAGAGAAAT	cjb111_all.seq							
201	AACAGTTGATAAACCAAAATACATTAAAAATTACGTTTAAACCAGAGAAAT	nem316_all.seq							
200	AACAGTTGATAAACCAAAATACATTAAAAATTACGTTTAAACCAGAGAAAT	a909_all.seq							
TTAAAGAAATTGCTGAGCTACTTAAAGCAATGACCCTTGTTAAAAATCAA Majority									
3260		3270		3280		3290		3300	
248	TTAAAGAAATTGCTGAGCTACTTAAAGCAATGACCCTTGTTAAAAATCAA	2603_all.seq							
217	TTAAAGAAATTGCTGAGCTACTTAAAGCAATGACCCTTGTTAAAAATCAA	18rs21_all.seq							
249	TTAAAGAAATTGCTGAGCTACTTAAAGCAATGACCCTTGTTAAAAATCAA	cohl1_all.seq							
249	TTAAAGAAATTGCTGAGCTACTTAAAGCAATGACCCTTGTTAAAAATCAA	cjb111_all.seq							
251	TTAAAGAAATTGCTGAGCTACTTAAAGCAATGACCCTTGTTAAAAATCAA	nem316_all.seq							
250	TTAAAGAAATTGCTGAGCTACTTAAAGCAATGACCCTTGTTAAAAATCAA	a909_all.seq							
GATGCTCTTGATAAAGCTACTGCAAATACAGATGATGCGGCATTTTTTTGA Majority									
3310		3320		3330		3340		3350	
298	GATGCTCTTGATAAAGCTACTGCAAATACAGATGATGCGGCATTTTTTTGA	2603_all.seq							
267	GATGCTCTTGATAAAGCTACTGCAAATACAGATGATGCGGCATTTTTTTGA	18rs21_all.seq							
299	GATGCTCTTGATAAAGCTACTGCAAATACAGATGATGCGGCATTTTTTTGA	cohl1_all.seq							
299	GATGCTCTTGATAAAGCTACTGCAAATACAGATGATGCGGCATTTTTTTGA	cjb111_all.seq							
301	GATGCTCTTGATAAAGCTACTGCAAATACAGATGATGCGGCATTTTTTTGA	nem316_all.seq							
300	GATGCTCTTGATAAAGCTACTGCAAATACAGATGATGCGGCATTTTTTTGA	a909_all.seq							
AATTCAGTTGCATCAACTATTAAATGAAAAAGCAGTTTTTAGGAAAAAGCAA Majority									
3360		3370		3380		3390		3400	
48	AATTCAGTTGCATCAACTATTAAATGAAAAAGCAGTTTTTAGGAAAAAGCAA	2603_all.seq							
17	AATTCAGTTGCATCAACTATTAAATGAAAAAGCAGTTTTTAGGAAAAAGCAA	18rs21_all.seq							
49	AATTCAGTTGCATCAACTATTAAATGAAAAAGCAGTTTTTAGGAAAAAGCAA	cohl1_all.seq							
49	AATTCAGTTGCATCAACTATTAAATGAAAAAGCAGTTTTTAGGAAAAAGCAA	cjb111_all.seq							
51	AATTCAGTTGCATCAACTATTAAATGAAAAAGCAGTTTTTAGGAAAAAGCAA	nem316_all.seq							
50	AATTCAGTTGCATCAACTATTAAATGAAAAAGCAGTTTTTAGGAAAAAGCAA	a909_all.seq							
TTGAAAAATACTTTTGAACCTTCAATATGACCATACTCCTGATAAAGCTGAC Majority									
3410		3420		3430		3440		3450	
98	TTGAAAAATACTTTTGAACCTTCAATATGACCATACTCCTGATAAAGCTGAC	2603_all.seq							
67	TTGAAAAATACTTTTGAACCTTCAATATGACCATACTCCTGATAAAGCTGAC	18rs21_all.seq							
99	TTGAAAAATACTTTTGAACCTTCAATATGACCATACTCCTGATAAAGCTGAC	cohl1_all.seq							
99	TTGAAAAATACTTTTGAACCTTCAATATGACCATACTCCTGATAAAGCTGAC	cjb111_all.seq							
01	TTGAAAAATACTTTTGAACCTTCAATATGACCATACTCCTGATAAAGCTGAC	nem316_all.seq							
00	TTGAAAAATACTTTTGAACCTTCAATATGACCATACTCCTGATAAAGCTGAC	a909_all.seq							
AATCCAAAAACCATCTAATCCTCCAAGAAAAACAGAAAGTTTCATACTGGTGG Majority									
3460		3470		3480		3490		3500	
48	AATCCAAAAACCATCTAATCCTCCAAGAAAAACAGAAAGTTTCATACTGGTGG	2603_all.seq							
17	AATCCAAAAACCATCTAATCCTCCAAGAAAAACAGAAAGTTTCATACTGGTGG	18rs21_all.seq							
49	AATCCAAAAACCATCTAATCCTCCAAGAAAAACAGAAAGTTTCATACTGGTGG	cohl1_all.seq							
49	AATCCAAAAACCATCTAATCCTCCAAGAAAAACAGAAAGTTTCATACTGGTGG	cjb111_all.seq							
51	AATCCAAAAACCATCTAATCCTCCAAGAAAAACAGAAAGTTTCATACTGGTGG	nem316_all.seq							
50	AATCCAAAAACCATCTAATCCTCCAAGAAAAACAGAAAGTTTCATACTGGTGG	a909_all.seq							

FIGURE 18 I

28/487

G A A A C C G A T T T G T A A A G A A A G A C T C A A C A A A A C A C T A G G T G G T G Majority											
		3510		3520		3530		3540		3550	
1498		G A A A C C A T T T T G T A A A G A A A G A C T C A A C A G A A A C A C A A A C A C T A G G T G G T G 2603_aill.seq									
1267		G A A A C C G A T T T T G T A A A G A A A G A C T C A A C A G A A A C A C A A A C A C T A G G T G G T G 18rs21_aill.seq									
1499		G A A A C C G A T T T T G T A A A G A A A G A C T C A A C A G A A A C A C A A A C A C T A G G T G G T G coh1_aill.seq									
1499		G A A A C C G A T T T T G T A A A G A A A G A C T C A A C A G A A A C A C A A A C A C T A G G T G G T G cjb111_aill.seq									
1501		G A A A C C G A T T T T G T A A A G A A A G A C T C A A C A G A A A C A C A A A C A C T A G G T G G T G nem316_aill.seq									
1500		G A A A C C G A T T T T G T A A A G A A A G A C T C A A C A G A A A C A C A A A C A C T A G G T G G T G a909_aill.seq									
C T G A G T T T G A T T T G T T G G C T T C T G A T G G G A C A G C A G T A A A A T G G A C A G A T Majority											
		3560		3570		3580		3590		3600	
1548		C T G A G T T T G A T T T G T T G G C T T C T G A T G G G A C A G C A G T A A A A T G G A C A G A T 2603_aill.seq									
1317		C T G A G T T T G A T T T G T T G G C T T C T G A T G G G A C A G C A G T A A A A T G G A C A G A T 18rs21_aill.seq									
1549		C T G A G T T T G A T T T G T T G G C T T C T G A T G G G A C A G C A G T A A A A T G G A C A G A T coh1_aill.seq									
1549		C T G A G T T T G A T T T G T T G G C T T C T G A T G G G A C A G C A G T A A A A T G G A C A G A T cjb111_aill.seq									
1551		C T G A G T T T G A T T T G T T G G C T T C T G A T G G G A C A G C A G T A A A A T G G A C A G A T nem316_aill.seq									
1550		C T G A G T T T G A T T T G T T G G C T T C T G A T G G G A C A G C A G T A A A A T G G A C A G A T a909_aill.seq									
G C T C T T A T T A A A G C C G A A T A C T A A T A A A A A C T A T A T T G C T G G A G A A G C T G T Majority											
		3610		3620		3630		3640		3650	
598		G C T C T T A T T A A A G C C G A A T A C T A A T A A A A A C T A T A T T G C T G G A G A A G C T G T 2603_aill.seq									
367		G C T C T T A T T A A A G C C G A A T A C T A A T A A A A A C T A T A T T G C T G G A G A A G C T G T 18rs21_aill.seq									
599		G C T C T T A T T A A A G C C G A A T A C T A A T A A A A A C T A T A T T G C T G G A G A A G C T G T coh1_aill.seq									
599		G C T C T T A T T A A A G C C G A A T A C T A A T A A A A A C T A T A T T G C T G G A G A A G C T G T cjb111_aill.seq									
601		G C T C T T A T T A A A G C C G A A T A C T A A T A A A A A C T A T A T T G C T G G A G A A G C T G T nem316_aill.seq									
600		G C T C T T A T T A A A G C C G A A T A C T A A T A A A A A C T A T A T T G C T G G A G A A G C T G T a909_aill.seq									
T A C T G G G C A A C C A A T C A A A T T G A A A T C A C A T A C A G A C G G T A C G T T T G A G A Majority											
		3660		3670		3680		3690		3700	
648		T A C T G G G C A A C C A A T C A A A T T G A A A T C A C A T A C A G A C G G T A C G T T T G A G A 2603_aill.seq									
417		T A C T G G G C A A C C A A T C A A A T T G A A A T C A C A T A C A G A C G G T A C G T T T G A G A 18rs21_aill.seq									
649		T A C T G G G C A A C C A A T C A A A T T G A A A T C A C A T A C A G A C G G T A C G T T T G A G A coh1_aill.seq									
649		T A C T G G G C A A C C A A T C A A A T T G A A A T C A C A T A C A G A C G G T A C G T T T G A G A cjb111_aill.seq									
651		T A C T G G G C A A C C A A T C A A A T T G A A A T C A C A T A C A G A C G G T A C G T T T G A G A nem316_aill.seq									
650		T A C T G G G C A A C C A A T C A A A T T G A A A T C A C A T A C A G A C G G T A C G T T T G A G A a909_aill.seq									
T T A A A G C T T T G C C T T A T G C A G T T G A T G C G A A T G C A G A G G G T A C A G C A G T A Majority											
		3710		3720		3730		3740		3750	
598		T T A A A G C T T T G C C T T A T G C A G T T G A T G C G A A T G C A G A G G G T A C A G C A G T A 2603_aill.seq									
467		T T A A A G C T T T G C C T T A T G C A G T T G A T G C G A A T G C A G A G G G T A C A G C A G T A 18rs21_aill.seq									
599		T T A A A G C T T T G C C T T A T G C A G T T G A T G C G A A T G C A G A G G G T A C A G C A G T A coh1_aill.seq									
599		T T A A A G C T T T G C C T T A T G C A G T T G A T G C G A A T G C A G A G G G T A C A G C A G T A cjb111_aill.seq									
701		T T A A A G C T T T G C C T T A T G C A G T T G A T G C G A A T G C A G A G G G T A C A G C A G T A nem316_aill.seq									
700		T T A A A G C T T T G C C T T A T G C A G T T G A T G C G A A T G C A G A G G G T A C A G C A G T A a909_aill.seq									
A C T T A C A A A T T A A A A G A A A C A A A A G C A C C A G A A G G T T A T G T A A T C C C T G A Majority											
		3760		3770		3780		3790		3800	
748		A C T T A C A A A T T A A A A G A A A C A A A A G C A C C A G A A G G T T A T G T A A T C C C T G A 2603_aill.seq									
517		A C T T A C A A A T T A A A A G A A A C A A A A G C A C C A G A A G G T T A T G T A A T C C C T G A 18rs21_aill.seq									
749		A C T T A C A A A T T A A A A G A A A C A A A A G C A C C A G A A G G T T A T G T A A T C C C T G A coh1_aill.seq									
749		A C T T A C A A A T T A A A A G A A A C A A A A G C A C C A G A A G G T T A T G T A A T C C C T G A cjb111_aill.seq									
751		A C T T A C A A A T T A A A A G A A A C A A A A G C A C C A G A A G G T T A T G T A A T C C C T G A nem316_aill.seq									
750		A C T T A C A A A T T A A A A G A A A C A A A A G C A C C A G A A G G T T A T G T A A T C C C T G A a909_aill.seq									
T A A A G A A A T C G A G T T T A C A G T A T C A C A A A C A T C T T A T A A A T A C A A A A C C A A Majority											
		3810		3820		3830		3840		3850	
798		T A A A G A A A T C G A G T T T A C A G T A T C A C A A A C A T C T T A T A A A T A C A A A A C C A A 2603_aill.seq									
567		T A A A G A A A T C G A G T T T A C A G T A T C A C A A A C A T C T T A T A A A T A C A A A A C C A A 18rs21_aill.seq									
799		T A A A G A A A T C G A G T T T A C A G T A T C A C A A A C A T C T T A T A A A T A C A A A A C C A A coh1_aill.seq									
799		T A A A G A A A T C G A G T T T A C A G T A T C A C A A A C A T C T T A T A A A T A C A A A A C C A A cjb111_aill.seq									
801		T A A A G A A A T C G A G T T T A C A G T A T C A C A A A C A T C T T A T A A A T A C A A A A C C A A nem316_aill.seq									
800		T A A A G A A A T C G A G T T T A C A G T A T C A C A A A C A T C T T A T A A A T A C A A A A C C A A a909_aill.seq									

FIGURE 18 J

FIGURE 18 K

30/487

T C A G T T G A C G A T T G T T C A T C T T G A A G C A A G G C A T A T T G A T C G T C C A A A T C Majority									
4210		4220		4230		4240		4250	
4198	T	C	A	G	T	T	G	A	C
3967	T	C	A	G	T	T	G	A	C
4199	T	C	A	G	T	T	G	A	C
4199	T	C	A	G	T	T	G	A	C
4201	T	C	A	G	T	T	G	A	C
4200	T	C	A	G	T	T	G	A	C
C A C A G T T G G A G A T T G C C C C T A A A G A A G G G A C T C C A A T T G A A G G A G T A C T C Majority									
4260		4270		4280		4290		4300	
4248	C	A	C	A	G	T	T	G	A
4017	C	A	C	A	G	T	T	G	A
4249	C	A	C	A	G	T	T	G	A
4249	C	A	C	A	G	T	T	G	A
4251	C	A	C	A	G	T	T	G	A
4250	C	A	C	A	G	T	T	G	A
T A T C A G T T G T A C C A A T T A A A A T C A A C T G A A G A T G G C G A T T T G T T G G C A C A Majority									
4310		4320		4330		4340		4350	
4298	T	A	T	C	A	G	T	T	G
1067	T	A	T	C	A	G	T	T	G
4299	T	A	T	C	A	G	T	T	G
4299	T	A	T	C	A	G	T	T	G
1301	T	A	T	C	A	G	T	T	G
1300	T	A	T	C	A	G	T	T	G
T T G G A A T T C C C T A A C T A T C A C A G A A T T G A A A A A C A G G C G C A G C A G G T T T Majority									
4360		4370		4380		4390		4400	
1348	T	T	G	G	A	A	T	T	C
1117	T	T	G	G	A	A	T	T	C
1349	T	T	G	G	A	A	T	T	C
1349	T	T	G	G	A	A	T	T	C
1351	T	T	G	G	A	A	T	T	C
1350	T	T	G	G	A	A	T	T	C
T T G A A G C C A C T A C T A A T C A A C A A G G A A A G G C T A C A T T T A A C C A A C T A C C A Majority									
4410		4420		4430		4440		4450	
1398	T	T	G	A	A	G	C	C	A
1167	T	T	G	A	A	G	C	C	A
1399	T	T	G	A	A	G	C	C	A
399	T	T	G	A	A	G	C	C	A
401	T	T	G	A	A	G	C	C	A
400	T	T	G	A	A	G	C	C	A
G A T G G A A T T T A T T A T G C T C T G G C G G T T A A A G C C G G T G A A A A A A A T C G T A A Majority									
4460		4470		4480		4490		4500	
448	G	A	T	G	G	A	A	T	T
217	G	A	T	G	G	A	A	T	T
449	G	A	T	G	G	A	A	T	T
449	G	A	T	G	G	A	A	T	T
451	G	A	T	G	G	A	A	T	T
450	G	A	T	G	G	A	A	T	T
T G T C T C A G C T T T C T T G G T T G A C T T G T C T G A G G A T A A A G T G A T T T A T C C T A Majority									
4510		4520		4530		4540		4550	
498	T	G	T	C	T	C	A	G	C
267	T	G	T	C	T	C	A	G	C
499	T	G	T	C	T	C	A	G	C
499	T	G	T	C	T	C	A	G	C
501	T	G	T	C	T	C	A	G	C
500	T	G	T	C	T	C	A	G	C

FIGURE 18 L

FIGURE 18 M

C C A T C T C G A G G A G G C T C T T A T T C C C A A A A C A G G T G A G C A A C A G G C C A A T G G C Majority											
		4910		4920		4930		4940		4950	
4898	C	C	A	T	C	T	C	G	A	G	G
4667	C	C	A	T	C	T	C	G	A	G	G
4899	C	C	A	T	C	T	C	G	A	G	G
4899	C	C	A	T	C	T	C	G	A	G	G
4901	C	C	A	T	C	T	C	G	A	G	G
4900	C	C	A	T	C	T	C	G	A	G	G
A C T T G T A A T T A T T G C T G G T A T T T T A A T T G C T T T A G C C T T A C G A T T A C T A T Majority											
		4960		4970		4980		4990		5000	
4948	A	C	T	T	G	T	A	A	T	T	A
4717	A	C	T	T	G	T	A	A	T	T	A
4949	A	C	T	T	G	T	A	A	T	T	A
4949	A	C	T	T	G	T	A	A	T	T	A
4951	A	C	T	T	G	T	A	A	T	T	A
4950	A	C	T	T	G	T	A	A	T	T	A
C A A A A C A T C G G A A A C A T C A A A A T A A G G A T T A G C A T G G G A C A A A A A T C A A A Majority											
		5010		5020		5030		5040		5050	
4998	C	A	A	A	A	C	A	T	C	G	A
4767	C	A	A	A	A	C	A	T	C	G	A
4999	C	A	A	A	A	C	A	T	C	G	A
4999	C	A	A	A	A	C	A	T	C	G	A
5001	C	A	A	A	A	C	A	T	C	G	A
5000	C	A	A	A	A	C	A	T	C	G	A
A A T A T C T C T A G C T A C G A A T A T T T C G T A T A T G G A T T T T T C G T T T A A T T T T C T Majority											
		5060		5070		5080		5090		5100	
5048	A	A	T	A	T	C	T	C	T	A	G
1817	A	A	T	A	T	C	T	C	T	A	G
5049	A	A	T	A	T	C	T	C	T	A	G
5049	A	A	T	A	T	C	T	C	T	A	G
5051	A	A	T	A	T	C	T	C	T	A	G
5050	A	A	T	A	T	C	T	C	T	A	G
T A G C G G G T T T C C T T G T T T T G G C A T T T C C C A T C G T T A G T C A G G T C A T G T A C Majority											
		5110		5120		5130		5140		5150	
5098	T	A	G	C	G	G	T	T	T	C	C
1867	T	A	G	C	G	G	T	T	T	C	C
5099	T	A	G	C	G	G	T	T	T	C	C
5099	T	A	G	C	G	G	T	T	T	C	C
5101	T	A	G	C	G	G	T	T	T	C	C
5100	T	A	G	C	G	G	T	T	T	C	C
T T T C A A G C C T C T C A C G C C A A T A T T A A T G C T T T T A A A G A A G C T G T T A C C A A Majority											
		5160		5170		5180		5190		5200	
5148	T	T	T	C	A	A	G	C	C	T	C
917	T	T	T	C	A	A	G	C	C	T	C
5149	T	T	T	C	A	A	G	C	C	T	C
5149	T	T	T	C	A	A	G	C	C	T	C
5151	T	T	T	C	A	A	G	C	C	T	C
5150	T	T	T	C	A	A	G	C	C	T	C
G A T T G A C C G G G T G G A G A T T A A T C G G C G T T T A G A A C T T G C T T A T G C T T A T A Majority											
		5210		5220		5230		5240		5250	
198	G	A	T	T	G	A	C	C	G	G	G
967	G	A	T	T	G	A	C	C	G	G	G
199	G	A	T	T	G	A	C	C	G	G	G
199	G	A	T	T	G	A	C	C	G	G	G
201	G	A	T	T	G	A	C	C	G	G	G
200	G	A	T	T	G	A	C	C	G	G	G

FIGURE 18 N

33/487

A C G C C A G T A T A G C A G G T G C C A A A A C T A A T G G C G A A T A T C C A G C G C T T A A A Majority																																																	
5260										5270										5280										5290										5300									
5248	A C G C C A G T A T A G C A G G T G C C A A A A C T A A T G G C G A A T A T C C A G C G C T T A A A																														2603_all.seq																		
5017	A C G C C A G T A T A G C A G G T G C C A A A A C T A A T G G C G A A T A T C C A G C G C T T A A A																														18rs21_all.seq																		
5249	A C G C C A G T A T A G C A G G T G C C A A A A C T A A T G G C G A A T A T C C A G C G C T T A A A																														cohl_all.seq																		
5249	A C G C C A G T A T A G C A G G T G C C A A A A C T A A T G G C G A A T A T C C A G C G C T T A A A																														cjb111_all.seq																		
5251	A C G C C A G T A T A G C A G G T G C C A A A A C T A A T G G C G A A T A T C C A G C G C T T A A A																														nea316_all.seq																		
5250	A C G C C A G T A T A G C A G G T G C C A A A A C T A A T G G C G A A T A T C C A G C G C T T A A A																														a909_all.seq																		
G A C C C C T A C T C T G C T G A A C A A A A G C A G G C A G G G T C G T T G A G T A C G C C C G Majority																																																	
5310										5320										5330										5340										5350									
5298	G A C C C C T A C T C T G C T G A A C A A A A G C A G G C A G G G G T C G T T G A G T A C G C C C G																														2603_all.seq																		
5067	G A C C C C T A C T C T G C T G A A C A A A A G C A G G C A G G G G T C G T T G A G T A C G C C C G																														18rs21_all.seq																		
5299	G A C C C C T A C T C T G C T G A A C A A A A G C A G G C A G G G G T C G T T G A G T A C G C C C G																														cohl_all.seq																		
5299	G A C C C C T A C T C T G C T G A A C A A A A G C A G G C A G G G G T C G T T G A G T A C G C C C G																														cjb111_all.seq																		
5301	G A C C C C T A C T C T G C T G A A C A A A A G C A G G C A G G G G T C G T T G A G T A C G C C C G																														nea316_all.seq																		
5300	G A C C C C T A C T C T G C T G A A C A A A A G C A G G C A G G G G T C G T T G A G T A C G C C C G																														a909_all.seq																		
C A T G C T T G A A G T C A A A G A A C A A A T A G G T C A T G T G A T T A T T C C A A G A A T T A Majority																																																	
5360										5370										5380										5390										5400									
5348	C A T G C T T G A A G T C A A A G A A C A A A T A G G T C A T G T G A T T A T T C C A A G A A T T A																														2603_all.seq																		
5117	C A T G C T T G A A G T C A A A G A A C A A A T A G G T C A T G T G A T T A T T C C A A G A A T T A																														18rs21_all.seq																		
5349	C A T G C T T G A A G T C A A A G A A C A A A T A G G T C A T G T G A T T A T T C C A A G A A T T A																														cohl_all.seq																		
5349	C A T G C T T G A A G T C A A A G A A C A A A T A G G T C A T G T G A T T A T T C C A A G A A T T A																														cjb111_all.seq																		
5351	C A T G C T T G A A G T C A A A G A A C A A A T A G G T C A T G T G A T T A T T C C A A G A A T T A																														nea316_all.seq																		
5350	C A T G C T T G A A G T C A A A G A A C A A A T A G G T C A T G T G A T T A T T C C A A G A A T T A																														a909_all.seq																		
A T C A G G A T A T C C C T A T T T A C G C T G G C T C T G C T G A A G A A A A T C T T C A G A G C Majority																																																	
5410										5420										5430										5440										5450									
398	A T C A G G A T A T C C C T A T T T A C G C T G G C T C T G C T G A A G A A A A T C T T C A G A G C																														2603_all.seq																		
167	A T C A G G A T A T C C C T A T T T A C G C T G G C T C T G C T G A A G A A A A T C T T C A G A G C																														18rs21_all.seq																		
399	A T C A G G A T A T C C C T A T T T A C G C T G G C T C T G C T G A A G A A A A T C T T C A G A G C																														cohl_all.seq																		
399	A T C A G G A T A T C C C T A T T T A C G C T G G C T C T G C T G A A G A A A A T C T T C A G A G C																														cjb111_all.seq																		
401	A T C A G G A T A T C C C T A T T T A C G C T G G C T C T G C T G A A G A A A A T C T T C A G A G C																														nea316_all.seq																		
400	A T C A G G A T A T C C C T A T T T A C G C T G G C T C T G C T G A A G A A A A T C T T C A G A G C																														a909_all.seq																		
G C C G T T G G A C A T T T A G A G G G G A C C A G T C T T C C A G T C G G T G G T G A G T C A A C Majority																																																	
5460										5470										5480										5490										5500									
448	G C C G T T G G A C A T T T A G A G G G G A C C A G T C T T T C C A G T C G G T G G T G A G T C A A C																														2603_all.seq																		
217	G C C G T T G G A C A T T T A G A G G G G A C C A G T C T T T C C A G T C G G T G G T G A G T C A A C																														18rs21_all.seq																		
449	G C C G T T G G A C A T T T A G A G G G G A C C A G T C T T T C C A G T C G G T G G T G A G T C A A C																														cohl_all.seq																		
449	G C C G T T G G A C A T T T A G A G G G G A C C A G T C T T T C C A G T C G G T G G T G A G T C A A C																														cjb111_all.seq																		
451	G C C G T T G G A C A T T T A G A G G G G A C C A G T C T T T C A G T C G G T G G T G A G T C A A C																														nea316_all.seq																		
450	G C C G T T G G A C A T T T A G A G G G G A C C A G T C T T T C A G T C G G T G G T G A G T C A A C																														a909_all.seq																		
T C A T G C C G T T C T A A C T G C C C A T C G A G G G C T A C C A A C G G C C A A G C T A T T T A Majority																																																	
5510										5520										5530										5540										5550									
498	T C A T G C C G T T C T A A C T G C C C A T C G A G G G C T A C C A A C G G C C A A G C T A T T T A																														2603_all.seq																		
267	T C A T G C C G T T C T A A C T G C C C A T C G A G G G C T A C C A A C G G C C A A G C T A T T T A																														18rs21_all.seq																		
499	T C A T G C C G T T C T A A C T G C C C A T C G A G G G C T A C C A A C G G C C A A G C T A T T T A																														cohl_all.seq																		
499	T C A T G C C G T T C T A A C T G C C C A T C G A G G G C T A C C A A C G G C C A A G C T A T T T A																														cjb111_all.seq																		
501	T C A T G C C G T T C T A A C T G C C C A T C G A G G G C T A C C A A C G G C C A A G C T A T T T A																														nea316_all.seq																		
500	T C A T G C C G T T C T A A C T G C C C A T C G A G G G C T A C C A A C G G C C A A G C T A T T T A																														a909_all.seq																		
C C A A T T T A G A C A A G G T A A C A G T A G G T G A C C G T T T T T A C A T T G A A C A C A T C Majority																																																	
5560										5570										5580										5590										5600									
548	C C A A T T T A G A C A A G G T A A C A G T A G G T G A C C G T T T T T A C A T T G A A C A C A T C																														2603_all.seq																		
517	C C A A T T T A G A C A A G G T A A C A G T A G G T G A C C G T T T T T A C A T T G A A C A C A T C																														18rs21_all.seq																		
549	C C A A T T T A G A C A A G G T A A C A G T A G G T G A C C G T T T T T A C A T T G A A C A C A T C																														cohl_all.seq																		
549	C C A A T T T A G A C A A G G T A A C A G T A G G T G A C C G T T T T T A C A T T G A A C A C A T C																														cjb111_all.seq																		
551	C C A A T T T A G A C A A G G T A A C A G T A G G T G A C C G T T T T T A C A T T G A A C A C A T C																														nea316_all.seq																		
550	C C A A T T T A G A C A A G G T A A C A G T A G G T G A C C G T T T T T A C A T T G A A C A C A T C																														a909_all.seq																		

Alignment Report of AI-1 alignment, using J: mem method with Weighted residue weight table.
Thursday, July 29, 2004 5:46 PM

FIGURE 18 P

FIGURE 18 Q

36/487

AAAGGCTAGCTGCGGCAATTTGAGGGAACCACTCTTCCAGTGGGAGGTTTGTCT Majority										
6310		6320		6330		6340		6350		
6298	AAAGGCTAGTGGGCCATTTGAGGGCAACCACTCTTCCAGTGGGAGGTTTGTCT									2603_all.seq
6067	AAAGGCTAGTGGGCCATTTGAGGGCAACCACTCTTCCAGTGGGAGGTTTGTCT									18rs21_all.seq
6299	AAAGGCTAGTGGGCCATTTGAGGGCAACCACTCTTCCAGTGGGAGGTTTGTCT									cohl_all.seq
6299	AAAGGCTAGTGGGCCATTTGAGGGCAACCACTCTTCCAGTGGGAGGTTTGTCT									cjb111_all.seq
6301	AAAGGCTAGTGGGCCATTTGAGGGCAACCACTCTTCCAGTGGGAGGTTTGTCT									nen316_all.seq
6300	AAAGGCTAGTGGGCCATTTGAGGGCAACCACTCTTCCAGTGGGAGGTTTGTCT									a909_all.seq
AACCATTTCAGTACTAACTGCCACCGTGGCTTGCCAAACAGCTAGGCTAT Majority										
6360		6370		6380		6390		6400		
6348	AACCATTTCAGTACTAACTGCCACCGTGGCTTGCCAAACAGCTAGGCTAT									2603_all.seq
6117	AACCATTTCAGTACTAACTGCCACCGTGGCTTGCCAAACAGCTAGGCTAT									18rs21_all.seq
6349	AACCATTTCAGTACTAACTGCCACCGTGGCTTGCCAAACAGCTAGGCTAT									cohl_all.seq
6349	AACCATTTCAGTACTAACTGCCACCGTGGCTTGCCAAACAGCTAGGCTAT									cjb111_all.seq
6351	AACCATTTCAGTACTAACTGCCACCGTGGCTTGCCAAACAGCTAGGCTAT									nen316_all.seq
6350	AACCATTTCAGTACTAACTGCCACCGTGGCTTGCCAAACAGCTAGGCTAT									a909_all.seq
TTACCGACTTAAATAAAAGTTAAAAAAGGCCACAGTTTTCTATGTGACGGAAC Majority										
6410		6420		6430		6440		6450		
6398	TTACCGACTTAAATAAAAGTTAAAAAAGGCCACAGTTTTCTATGTGACGGAAC									2603_all.seq
6167	TTACCGACTTAAATAAAAGTTAAAAAAGGCCACAGTTTTCTATGTGACGGAAC									18rs21_all.seq
6399	TTACCGACTTAAATAAAAGTTAAAAAAGGCCACAGTTTTCTATGTGACGGAAC									cohl_all.seq
6399	TTACCGACTTAAATAAAAGTTAAAAAAGGCCACAGTTTTCTATGTGACGGAAC									cjb111_all.seq
6401	TTACCGACTTAAATAAAAGTTAAAAAAGGCCACAGTTTTCTATGTGACGGAAC									nen316_all.seq
6400	TTACCGACTTAAATAAAAGTTAAAAAAGGCCACAGTTTTCTATGTGACGGAAC									a909_all.seq
ATCAAGGAAACACTTGCCTACAAAGTCGTCTCTATCAAAGTTCTGGATCC Majority										
6460		6470		6480		6490		6500		
6448	ATCAAGGAAACACTTGCCTACAAAGTCGTCTCTATCAAAGTTCTGGATCC									2603_all.seq
6217	ATCAAGGAAACACTTGCCTACAAAGTCGTCTCTATCAAAGTTCTGGATCC									18rs21_all.seq
6449	ATCAAGGAAACACTTGCCTACAAAGTCGTCTCTATCAAAGTTCTGGATCC									cohl_all.seq
6449	ATCAAGGAAACACTTGCCTACAAAGTCGTCTCTATCAAAGTTCTGGATCC									cjb111_all.seq
6451	ATCAAGGAAACACTTGCCTACAAAGTCGTCTCTATCAAAGTTCTGGATCC									nen316_all.seq
6450	ATCAAGGAAACACTTGCCTACAAAGTCGTCTCTATCAAAGTTCTGGATCC									a909_all.seq
AACAGCTTTTAAAGTGAGGTTAAGATTGCTCAATGCTAAGGATTATATAACCT Majority										
6510		6520		6530		6540		6550		
3498	AACAGCTTTTAAAGTGAGGTTAAGATTGCTCAATGCTAAGGATTATATAACCT									2603_all.seq
3267	AACAGCTTTTAAAGTGAGGTTAAGATTGCTCAATGCTAAGGATTATATAACCT									18rs21_all.seq
3499	AACAGCTTTTAAAGTGAGGTTAAGATTGCTCAATGCTAAGGATTATATAACCT									cohl_all.seq
3499	AACAGCTTTTAAAGTGAGGTTAAGATTGCTCAATGCTAAGGATTATATAACCT									cjb111_all.seq
3501	AACAGCTTTTAAAGTGAGGTTAAGATTGCTCAATGCTAAGGATTATATAACCT									nen316_all.seq
3500	AACAGCTTTTAAAGTGAGGTTAAGATTGCTCAATGCTAAGGATTATATAACCT									a909_all.seq
TGCTGACTTGCACACCTTACATCATCAATAGTCATCGTCTCTTGGTAAAAA Majority										
6560		6570		6580		6590		6600		
1548	TGCTGACTTGCACACCTTACATCATCAATAGTCATCGTCTCTTGGTAAAAA									2603_all.seq
1317	TGCTGACTTGCACACCTTACATCATCAATAGTCATCGTCTCTTGGTAAAAA									18rs21_all.seq
1549	TGCTGACTTGCACACCTTACATCATCAATAGTCATCGTCTCTTGGTAAAAA									cohl_all.seq
1549	TGCTGACTTGCACACCTTACATCATCAATAGTCATCGTCTCTTGGTAAAAA									cjb111_all.seq
1551	TGCTGACTTGCACACCTTACATCATCAATAGTCATCGTCTCTTGGTAAAAA									nen316_all.seq
1550	TGCTGACTTGCACACCTTACATCATCAATAGTCATCGTCTCTTGGTAAAAA									a909_all.seq
CGAGAGCGCTATTGCTTATGATTCTACCGAGGCGGAAAGGCACAAAGAACAA Majority										
6610		6620		6630		6640		6650		
1598	CGAGAGCGCTATTGCTTATGATTCTACCGAGGCGGAAAGGCACAAAGAACAA									2603_all.seq
1367	CGAGAGCGCTATTGCTTATGATTCTACCGAGGCGGAAAGGCACAAAGAACAA									18rs21_all.seq
1599	CGAGAGCGCTATTGCTTATGATTCTACCGAGGCGGAAAGGCACAAAGAACAA									cohl_all.seq
1599	CGAGAGCGCTATTGCTTATGATTCTACCGAGGCGGAAAGGCACAAAGAACAA									cjb111_all.seq
1601	CGAGAGCGCTATTGCTTATGATTCTACCGAGGCGGAAAGGCACAAAGAACAA									nen316_all.seq
1600	CGAGAGCGCTATTGCTTATGATTCTACCGAGGCGGAAAGGCACAAAGAACAA									a909_all.seq

FIGURE 18 R

37/487

A A C C G T A C A A G A T T A T C G T T T G T C A C T A G T G T T G A A G A T A C T A C T A G T A T Majority									
6660		6670		6680		6690		6700	
1648	A A C C G T A C A A G A T T A T C G T T T G T C A C T A G T G T T G A A G A T A C T A C T A G T A T	2603_all.seq							
1417	A A C C G T A C A A G A T T A T C G T T T G T C A C T A G T G T T G A A G A T A C T A C T A G T A T	18rs21_all.seq							
1649	A A C C G T A C A A G A T T A T C G T T T G T C A C T A G T G T T G A A G A T A C T A C T A G T A T	cohl_all.seq							
1649	A A C C G T A C A A G A T T A T C G T T T G T C A C T A G T G T T G A A G A T A C T A C T A G T A T	cjb111_all.seq							
1651	A A C C G T A C A A G A T T A T C G T T T G T C A C T A G T G T T G A A G A T A C T A C T A G T A T	nem316_all.seq							
1650	A A C C G T A C A A G A T T A T C G T T T G T C A C T A G T G T T G A A G A T A C T A C T A G T A T	a909_all.seq							
T A T T A A T T G G A C T C T T C A T C G T G A T A A T G A T G A G A A G A T G G A T G C A A C A T Majority									
6710		6720		6730		6740		6750	
1698	T A T T A A T T G G A C T C T T C A T C G T G A T A A T G A T G A G A A G A T G G A T G C A A C A T	2603_all.seq							
1467	T A T T A A T T G G A C T C T T C A T C G T G A T A A T G A T G A G A A G A T G G A T G C A A C A T	18rs21_all.seq							
1699	T A T T A A T T G G A C T C T T C A T C G T G A T A A T G A T G A G A A G A T G G A T G C A A C A T	cohl_all.seq							
1699	T A T T A A T T G G A C T C T T C A T C G T G A T A A T G A T G A G A A G A T G G A T G C A A C A T	cjb111_all.seq							
1701	T A T T A A T T G G A C T C T T C A T C G T G A T A A T G A T G A G A A G A T G G A T G C A A C A T	nem316_all.seq							
1700	T A T T A A T T G G A C T C T T C A T C G T G A T A A T G A T G A G A A G A T G G A T G C A A C A T	a909_all.seq							
C G T C A A T A A C G A T G T T G T G A A T G G C T T A C T T A C T T A T C A A A T A G G T G A C T Majority									
6760		6770		6780		6790		6800	
1748	C G T C A A T A A C G A T G T T G T G A A T G G C T T A C T T A C T T A T C A A A T A G G T G A C T	2603_all.seq							
517	C G T C A A T A A C G A T G T T G T G A A T G G C T T A C T T A C T T A T C A A A T A G G T G A C T	18rs21_all.seq							
1749	C G T C A A T A A C G A T G T T G T G A A T G G C T T A C T T A C T T A T C A A A T A G G T G A C T	cohl_all.seq							
1749	C G T C A A T A A C G A T G T T G T G A A T G G C T T A C T T A C T T A T C A A A T A G G T G A C T	cjb111_all.seq							
1751	C G T C A A T A A C G A T G T T G T G A A T G G C T T A C T T A C T T A T C A A A T A G G T G A C T	nem316_all.seq							
1750	C G T C A A T A A C G A T G T T G T G A A T G G C T T A C T T A C T T A T C A A A T A G G T G A C T	a909_all.seq							
A A T G A T G A T T G T G A A T A A T G G T T A T C T A G A A G G C A G A A A A A T G A A A A A G A Majority									
6810		6820		6830		6840		6850	
1794	A A T G A T G A T T G T G A A T A A T G G T T A T C T A G A A G G C A G A A A A A T G A A A A A G A	2603_all.seq							
567	A A T G A T G A T T G T G A A T A A T G G T T A T C T A G A A G G C A G A A A A A T G A A A A A G A	18rs21_all.seq							
1799	A A T G A T G A T T G T G A A T A A T G G T T A T C T A G A A G G C A G A A A A A T G A A A A A G A	cohl_all.seq							
1799	A A T G A T G A T T G T G A A T A A T G G T T A T C T A G A A G G C A G A A A A A T G A A A A A G A	cjb111_all.seq							
801	A A T G A T G A T T G T G A A T A A T G G T T A T C T A G A A G G C A G A A A A A T G A A A A A G A	nem316_all.seq							
800	A A T G A T G A T T G T G A A T A A T G G T T A T C T A G A A G G C A G A A A A A T G A A A A A G A	a909_all.seq							
G A C A A A A A A T A T G G A G A G G G T T A T C A G T T A C T T T A C T A A T C C T G T C C C A A Majority									
6860		6870		6880		6890		6900	
1844	G A C A A A A A A T A T G G A G A G G G T T A T C A G T T A C T T T A C T A A T C C T G T C C C A A	2603_all.seq							
1817	G A C A A A A A A T A T G G A G A G G G T T A T C A G T T A C T T T A C T A A T C C T G T C C C A A	18rs21_all.seq							
1849	G A C A A A A A A T A T G G A G A G G G T T A T C A G T T A C T T T A C T A A T C C T G T C C C A A	cohl_all.seq							
1849	G A C A A A A A A T A T G G A G A G G G T T A T C A G T T A C T T T A C T A A T C C T G T C C C A A	cjb111_all.seq							
1851	G A C A A A A A A T A T G G A G A G G G T T A T C A G T T A C T T T A C T A A T C C T G T C C C A A	nem316_all.seq							
1850	G A C A A A A A A T A T G G A G A G G G T T A T C A G T T A C T T T A C T A A T C C T G T C C C A A	a909_all.seq							
A T T C C A T T T G G T A T A T T G G T A C A A G G T G A A A C C C A A G A T A C C A A T C A A G C Majority									
6910		6920		6930		6940		6950	
1894	A T T C C A T T T G G T A T A T T G G T A C A A G G T G A A A C C C A A G A T A C C A A T C A A G C	2603_all.seq							
1867	A T T C C A T T T G G T A T A T T G G T A C A A G G T G A A A C C C A A G A T A C C A A T C A A G C	18rs21_all.seq							
1899	A T T C C A T T T G G T A T A T T G G T A C A A G G T G A A A C C C A A G A T A C C A A T C A A G C	cohl_all.seq							
1899	A T T C C A T T T G G T A T A T T G G T A C A A G G T G A A A C C C A A G A T A C C A A T C A A G C	cjb111_all.seq							
1901	A T T C C A T T T G G T A T A T T G G T A C A A G G T G A A A C C C A A G A T A C C A A T C A A G C	nem316_all.seq							
1900	A T T C C A T T T G G T A T A T T G G T A C A A G G T G A A A C C C A A G A T A C C A A T C A A G C	a909_all.seq							
A C T T G G A A A A G T A A T T G T T A A A A A A A C G G G A G A C A A T G C T A C A C C A T T A G Majority									
6960		6970		6980		6990		7000	
1944	A C T T G G A A A A G T A A T T G T T A A A A A A A C G G G A G A C A A T G C T A C A C C A T T A G	2603_all.seq							
1917	A C T T G G A A A A G T A A T T G T T A A A A A A A C G G G A G A C A A T G C T A C A C C A T T A G	18rs21_all.seq							
1949	A C T T G G A A A A G T A A T T G T T A A A A A A A C G G G A G A C A A T G C T A C A C C A T T A G	cohl_all.seq							
1949	A C T T G G A A A A G T A A T T G T T A A A A A A A C G G G A G A C A A T G C T A C A C C A T T A G	cjb111_all.seq							
1951	A C T T G G A A A A G T A A T T G T T A A A A A A A C G G G A G A C A A T G C T A C A C C A T T A G	nem316_all.seq							
1950	A C T T G G A A A A G T A A T T G T T A A A A A A A C G G G A G A C A A T G C T A C A C C A T T A G	a909_all.seq							

FIGURE 18 S

38/487

WO 2006/078318

Alignment Report of A1-1 alignment, using J. Hein method with Weighted residue weight table.
Thursday, July 29, 2004 5:46 PM

PCT/US2005/027239

Page 21

G C A A A G C G A C T T T T G T G T T A A A A A A T G A C A A T G A T A A G T C A G A A A C A A G T Majority									
7010		7020		7030		7040		7050	
5994	G	C	A	A	A	G	C	G	A
5767	G	C	A	A	A	G	C	G	A
5999	G	C	A	A	A	G	C	G	A
5999	G	C	A	A	A	G	C	G	A
7001	G	C	A	A	A	G	C	G	A
7000	G	C	A	A	A	G	C	G	A
C A C G A A A C G G T A G A G G G T T C T G G A G A A G C A A C C T T T G A A A A C A T A A A A C C Majority									
7060		7070		7080		7090		7100	
7044	C	A	C	G	A	A	A	C	G
5817	C	A	C	G	A	A	A	C	G
7049	C	A	C	G	A	A	A	C	G
7049	C	A	C	G	A	A	A	C	G
7051	C	A	C	G	A	A	A	C	G
7050	C	A	C	G	A	A	A	C	G
T G G A G A C T A C A C A T T A A G A G A A G A A A C A G C A C C A A T T G G T T A T A A A A A A A Majority									
7110		7120		7130		7140		7150	
7094	T	G	G	A	G	A	C	T	A
5867	T	G	G	A	G	A	C	T	A
7099	T	G	G	A	G	A	C	T	A
7099	T	G	G	A	G	A	C	T	A
7101	T	G	G	A	G	A	C	T	A
7100	T	G	G	A	G	A	C	T	A
C T G A T A A A A C C T G G A A A G T T A A A G T T G C A G A T A A C G G A G C A A C A A T A A T C Majority									
7160		7170		7180		7190		7200	
7144	C	T	G	A	T	A	A	A	C
5917	C	T	G	A	T	A	A	A	C
7149	C	T	G	A	T	A	A	A	C
7149	C	T	G	A	T	A	A	A	C
7151	C	T	G	A	T	A	A	A	C
7150	C	T	G	A	T	A	A	A	C
G A G G G T A T G G A T G C A G A T A A A G C A G A G A A A C G A A A A G A A G T T T T C A A T G C Majority									
7210		7220		7230		7240		7250	
7194	G	A	G	G	T	A	T	G	G
5967	G	A	G	G	T	A	T	G	G
7199	G	A	G	G	T	A	T	G	G
7199	G	A	G	G	T	A	T	G	G
7201	G	A	G	G	T	A	T	G	G
7200	G	A	G	G	T	A	T	G	G
C C A A T A T C C A A A A T C A G C T A T T T A T G A G G A T A C A A A A G A A A A T T A C C C A T Majority									
7260		7270		7280		7290		7300	
7244	C	C	A	A	T	A	T	C	C
7017	C	C	A	A	T	A	T	C	C
7249	C	C	A	A	T	A	T	C	C
7249	C	C	A	A	T	A	T	C	C
7251	C	C	A	A	T	A	T	C	C
7250	C	C	A	A	T	A	T	C	C
T A G T T A A T G T A G A G G G T T C C A A A G T T G G T G A A C A A T A C A A A G C A T T G A A T Majority									
7310		7320		7330		7340		7350	
7294	T	A	G	T	T	A	A	T	G
7067	T	A	G	T	T	A	A	T	G
7299	T	A	G	T	T	A	A	T	G
7299	T	A	G	T	T	A	A	T	G
7301	T	A	G	T	T	A	A	T	G
7300	T	A	G	T	T	A	A	T	G

FIGURE 18T

	CCAATAAATGCAAAAGATGCTCGAAGAGAGATTGCTGAAGCTTGCTTATC Majority									
	7360		7370		7380		7390		7400	
7344	CCAATAAATGCAAAAGATGCTCGAAGAGAGATTGCTGAAGCTTGCTTATC 2603_all.seq									
7117	CCAATAAATGCAAAAGATGCTCGAAGAGAGATTGCTGAAGCTTGCTTATC 18rs21_all.seq									
7349	CCAATAAATGCAAAAGATGCTCGAAGAGAGATTGCTGAAGCTTGCTTATC coh1_all.seq									
7349	CCAATAAATGCAAAAGATGCTCGAAGAGAGATTGCTGAAGCTTGCTTATC cjb111_all.seq									
7351	CCAATAAATGCAAAAGATGCTCGAAGAGAGATTGCTGAAGCTTGCTTATC nem316_all.seq									
7350	CCAATAAATGCAAAAGATGCTCGAAGAGAGATTGCTGAAGCTTGCTTATC a909_all.seq									
	AAAAAAAATTACAGGGGTCAATGATCTCGATAAGCAATAAATATAAAAATTG Majority									
	7410		7420		7430		7440		7450	
7394	AAAAAAAATTACAGGGGTCAATGATCTCGATAAGCAATAAATATAAAAATTG 2603_all.seq									
167	AAAAAAAATTACAGGGGTCAATGATCTCGATAAGCAATAAATATAAAAATTG 18rs21_all.seq									
7399	AAAAAAAATTACAGGGGTCAATGATCTCGATAAGCAATAAATATAAAAATTG coh1_all.seq									
7399	AAAAAAAATTACAGGGGTCAATGATCTCGATAAGCAATAAATATAAAAATTG cjb111_all.seq									
7401	AAAAAAAATTACAGGGGTCAATGATCTCGATAAGCAATAAATATAAAAATTG nem316_all.seq									
7400	AAAAAAAATTACAGGGGTCAATGATCTCGATAAGCAATAAATATAAAAATTG a909_all.seq									
	AATTAACCTGTTGAGGGTAAACCACCTGTTGAAACGAAAGAACTTAATCAA Majority									
	7460		7470		7480		7490		7500	
7444	AATTAACCTGTTGAGGGTAAACCACCTGTTGAAACGAAAGAACTTAATCAA 2603_all.seq									
217	AATTAACCTGTTGAGGGTAAACCACCTGTTGAAACGAAAGAACTTAATCAA 18rs21_all.seq									
7449	AATTAACCTGTTGAGGGTAAACCACCTGTTGAAACGAAAGAACTTAATCAA coh1_all.seq									
7449	AATTAACCTGTTGAGGGTAAACCACCTGTTGAAACGAAAGAACTTAATCAA cjb111_all.seq									
7451	AATTAACCTGTTGAGGGTAAACCACCTGTTGAAACGAAAGAACTTAATCAA nem316_all.seq									
7450	AATTAACCTGTTGAGGGTAAACCACCTGTTGAAACGAAAGAACTTAATCAA a909_all.seq									
	CCACTAGATGTCGTTGCTGCTATTAGATAAATTCAAATAGTATGAATAAATGA Majority									
	7510		7520		7530		7540		7550	
7494	CCACTAGATGTCGTTGCTGCTATTAGATAAATTCAAATAGTATGAATAAATGA 2603_all.seq									
267	CCACTAGATGTCGTTGCTGCTATTAGATAAATTCAAATAGTATGAATAAATGA 18rs21_all.seq									
7499	CCACTAGATGTCGTTGCTGCTATTAGATAAATTCAAATAGTATGAATAAATGA coh1_all.seq									
7499	CCACTAGATGTCGTTGCTGCTATTAGATAAATTCAAATAGTATGAATAAATGA cjb111_all.seq									
501	CCACTAGATGTCGTTGCTGCTATTAGATAAATTCAAATAGTATGAATAAATGA nem316_all.seq									
500	CCACTAGATGTCGTTGCTGCTATTAGATAAATTCAAATAGTATGAATAAATGA a909_all.seq									
	AAGAGCCCAATAATTCTCAAAGAGCATTAAAGCTGGGGAAGCAGTTGAAA Majority									
	7560		7570		7580		7590		7600	
544	AAGAGCCCAATAATTCTCAAAGAGCATTAAAGCTGGGGAAGCAGTTGAAA 2603_all.seq									
317	AAGAGCCCAATAATTCTCAAAGAGCATTAAAGCTGGGGAAGCAGTTGAAA 18rs21_all.seq									
549	AAGAGCCCAATAATTCTCAAAGAGCATTAAAGCTGGGGAAGCAGTTGAAA coh1_all.seq									
549	AAGAGCCCAATAATTCTCAAAGAGCATTAAAGCTGGGGAAGCAGTTGAAA cjb111_all.seq									
551	AAGAGCCCAATAATTCTCAAAGAGCATTAAAGCTGGGGAAGCAGTTGAAA nem316_all.seq									
550	AAGAGCCCAATAATTCTCAAAGAGCATTAAAGCTGGGGAAGCAGTTGAAA a909_all.seq									
	AGCTGATTGATAAAAATTACATCAAATAAAGACAATAAGCTAGCTCTTCTG Majority									
	7610		7620		7630		7640		7650	
594	AGCTGATTGATAAAAATTACATCAAATAAAGACAATAAGCTAGCTCTTCTG 2603_all.seq									
367	AGCTGATTGATAAAAATTACATCAAATAAAGACAATAAGCTAGCTCTTCTG 18rs21_all.seq									
599	AGCTGATTGATAAAAATTACATCAAATAAAGACAATAAGCTAGCTCTTCTG coh1_all.seq									
599	AGCTGATTGATAAAAATTACATCAAATAAAGACAATAAGCTAGCTCTTCTG cjb111_all.seq									
601	AGCTGATTGATAAAAATTACATCAAATAAAGACAATAAGCTAGCTCTTCTG nem316_all.seq									
600	AGCTGATTGATAAAAATTACATCAAATAAAGACAATAAGCTAGCTCTTCTG a909_all.seq									
	ACATATGCCCTCAACCATTGATGCTACTGAAGCGACCGTATCAAAGCG Majority									
	7660		7670		7680		7690		7700	
644	ACATATGCCCTCAACCATTGATGCTACTGAAGCGACCGTATCAAAGCG 2603_all.seq									
417	ACATATGCCCTCAACCATTGATGCTACTGAAGCGACCGTATCAAAGCG 18rs21_all.seq									
649	ACATATGCCCTCAACCATTGATGCTACTGAAGCGACCGTATCAAAGCG coh1_all.seq									
649	ACATATGCCCTCAACCATTGATGCTACTGAAGCGACCGTATCAAAGCG cjb111_all.seq									
651	ACATATGCCCTCAACCATTGATGCTACTGAAGCGACCGTATCAAAGCG nem316_all.seq									
650	ACATATGCCCTCAACCATTGATGCTACTGAAGCGACCGTATCAAAGCG a909_all.seq									

FIGURE 18 U

A C T T G C C G A T C A A A A T G G T A A A G C G C T G A A T G A T A G T G T A T C A T G G G A T T Majority					
	7710	7720	7730	7740	7750
1694	A C T T G C C G A T C A A A A T G G T A A A G C G C T G A A T G A T A G T G T A T C A T G G G A T T				2603_all.seq
1467	A G T T G C C G A T C A A A A T G G T A A A G C G C T G A A T G A T A G T G T A T C A T G G G A T T				18rs21_all.seq
1699	A G T T G C C G A T C A A A A T G G T A A A G C G C T G A A T G A T A G T G T A T C A T G G G A T T				cohl_all.seq
1699	A C T T G C C G A T C A A A A T G G T A A A G C G C T G A A T G A T A G T G T A T C A T G G G A T T				cjb111_all.seq
1701	A G T T G C C G A T C A A A A T G G T A A A G C G C T G A A T G A T A G T G T A T C A T G G G A T T				nea316_all.seq
1700	A G T T G C C G A T C A A A A T G G T A A A G C G C T G A A T G A T A G T G T A T C A T G G G A T T				a909_all.seq
A T C A T A A A A C T A C T T T T T A C A G C A A C T A C A C A T A A T T A C A G T T A T T T A A A T Majority					
	7760	7770	7780	7790	7800
744	A T C A T A A A A C T A C T T T T T A C A G C A A C T A C A C A T A A T T A C A G T T A T T T A A A T				2603_all.seq
517	A T C A T A A A A C T A C T T T T T A C A G C A A C T A C A C A T A A T T A C A C T T A T T T A A A T				18rs21_all.seq
749	A T C A T A A A A C T A C T T T T T A C A G C A A C T A C A C A T A A T T A C A C T T A T T T A A A T				cohl_all.seq
749	A T C A T A A A A C T A C T T T T T A C A G C A A C T A C A C A T A A T T A C A G T T A T T T A A A T				cjb111_all.seq
751	A T C A T A A A A C T A C T T T T T A C A G C A A C T A C A C A T A A T T A C A G T T A T T T A A A T				nea316_all.seq
750	A T C A T A A A A C T A C T T T T T A C A G C A A C T A C A C A T A A T T A C A G T T A T T T A A A T				a909_all.seq
T T A A C A A A T G A T G C T A A C G A A G T T A A T A T T C T A A A G T C A A G A A T T C C A A A Majority					
	7810	7820	7830	7840	7850
794	T T A A C A A A T G A T G C T A A C G A A G T T A A T A T T C T A A A G T C A A G A A T T C C A A A				2603_all.seq
567	T T A A C A A A T G A T G C T A A C G A A G T T A A T A T T C T A A A G T C A A G A A T T C C A A A				18rs21_all.seq
799	T T A A C A A A T G A T G C T A A C G A A G T T A A T A T T C T A A A G T C A A G A A T T C C A A A				cohl_all.seq
799	T T A A C A A A T G A T G C T A A C G A A G T T A A T A T T C T A A A G T C A A G A A T T C C A A A				cjb111_all.seq
801	T T A A C A A A T G A T G C T A A C G A A G T T A A T A T T C T A A A G T C A A G A A T T C C A A A				nea316_all.seq
800	T T A A C A A A T G A T G C T A A C G A A G T T A A T A T T C T A A A G T C A A G A A T T C C A A A				a909_all.seq
G G A A G C G G A G C A T A T A A A T G G G G A T C G C A C G C T C T A T C A A T T T G G T G C G A Majority					
	7860	7870	7880	7890	7900
844	G G A A G C G G A G C A T A T A A A T G G G G A T C G C A C G C T C T A T C A A T T T G G T G C G A				2603_all.seq
617	G G A A G C G G A G C A T A T A A A T G G G G A T C G C A C G C T C T A T C A A T T T G G T G C G A				18rs21_all.seq
849	G G A A G C G G A G C A T A T A A A T G G G G A T C G C A C G C T C T A T C A A T T T G G T G C G A				cohl_all.seq
849	G G A A G C G G A G C A T A T A A A T G G G G A T C G C A C G C T C T A T C A A T T T G G T G C G A				cjb111_all.seq
851	G G A A G C G G A G C A T A T A A A T G G G G A T C G C A C G C T C T A T C A A T T T G G T G C G A				nea316_all.seq
850	G G A A G C G G A G C A T A T A A A T G G G G A T C G C A C G C T C T A T C A A T T T G G T G C G A				a909_all.seq
C A T T T A C T C A A A A A G C T C T A A T G A A A G C A A A T G A A A T T T T A G A G A C A C A A Majority					
	7910	7920	7930	7940	7950
894	C A T T T A C T C A A A A A G C T C T A A T G A A A G C A A A T G A A A T T T T A G A G A C A C A A				2603_all.seq
667	C A T T T A C T C A A A A A G C T C T A A T G A A A G C A A A T G A A A T T T T A G A G A C A C A A				18rs21_all.seq
899	C A T T T A C T C A A A A A G C T C T A A T G A A A G C A A A T G A A A T T T T A G A G A C A C A A				cohl_all.seq
899	C A T T T A C T C A A A A A G C T C T A A T G A A A G C A A A T G A A A T T T T A G A G A C A C A A				cjb111_all.seq
901	C A T T T A C T C A A A A A G C T C T A A T G A A A G C A A A T G A A A T T T T A G A G A C A C A A				nea316_all.seq
900	C A T T T A C T C A A A A A G C T C T A A T G A A A G C A A A T G A A A T T T T A G A G A C A C A A				a909_all.seq
A G T T C T A A T G C T A G A A A A A A C T T A T T T T T C A C G T A A C T G A T G G T G T C C C Majority					
	7960	7970	7980	7990	8000
944	A G T T C T A A T G C T A G A A A A A A C T T A T T T T T C A C G T A A C T G A T G G T G T C C C				2603_all.seq
717	A G T T C T A A T G C T A G A A A A A A C T T A T T T T T C A C G T A A C T G A T G G T G T C C C				18rs21_all.seq
949	A G T T C T A A T G C T A G A A A A A A C T T A T T T T T C A C G T A A C T G A T G G T G T C C C				cohl_all.seq
949	A G T T C T A A T G C T A G A A A A A A C T T A T T T T T C A C G T A A C T G A T G G T G T C C C				cjb111_all.seq
951	A G T T C T A A T G C T A G A A A A A A C T T A T T T T T C A C G T A A C T G A T G G T G T C C C				nea316_all.seq
950	A G T T C T A A T G C T A G A A A A A A C T T A T T T T T C A C G T A A C T G A T G G T G T C C C				a909_all.seq
T A C G A T G T C T T A T G C C A T A A A T T T T A A T C C T T A T A T A T C A A C A T C T T A C C Majority					
	8010	8020	8030	8040	8050
994	T A C G A T G T C T T A T G C C A T A A A T T T T A A T C C T T A T A T A T A T C A A C A T C T T A C C				2603_all.seq
767	T A C G A T G T C T T A T G C C A T A A A T T T T A A T C C T T A T A T A T A T C A A C A T C T T A C C				18rs21_all.seq
999	T A C G A T G T C T T A T G C C A T A A A T T T T A A T C C T T A T A T A T C A A C A T C T T A C C				cohl_all.seq
999	T A C G A T G T C T T A T G C C A T A A A T T T T A A T C C T T A T A T A T A T C A A C A T C T T A C C				cjb111_all.seq
301	T A C G A T G T C T T A T G C C A T A A A T T T T A A T C C T T A T A T A T A T C A A C A T C T T A C C				nea316_all.seq
300	T A C G A T G T C T T A T G C C A T A A A T T T T A A T C C T T A T A T A T A T C A A C A T C T T A C C				a909_all.seq

FIGURE 18 V

41/487

WO 2006/078318

Alignment Report of A1-1 alignment, using J. Hein method with Weighted residue weight table.
Thursday, July 29, 2004 5:46 PM

PCT/US2005/027239

Page 24

A A A A C C A G C T T T A A T T C T T T T T T A A A T A A A A T A C C A G A T A G A A G T C G T A T T Majority											
		8060		8070		8080		8090		8100	
1044		A A A A C C A G T T T A A T T C T T T T T T A A A T A A A A T A C C A G A T A G A A G T C G T A T T 2603_all.seq									
1817		A A A A C C A G T T T A A T T C T T T T T T A A A T A A A A T A C C A G A T A G A A G T C G T A T T 18rs21_all.seq									
1049		A A A A C C A G T T T A A T T C T T T T T T A A A T A A A A T A C C A G A T A G A A G T C G T A T T coh1_all.seq									
1051		A A A A C C A G T T T A A T T C T T T T T T A A A T A A A A T A C C A G A T A G A A G T C G T A T T cjb111_all.seq									
1050		A A A A C C A G T T T A A T T C T T T T T T A A A T A A A A T A C C A G A T A G A A G T C G T A T T nem316_all.seq									
		A A A A C C A G T T T A A T T C T T T T T T A A A T A A A A T A C C A G A T A G A A G T C G T A T T a909_all.seq									
C T C C A A G A G C A T T T T A T A A T C A A T G C T G A T G A T T A T C A A A T A G T A A A A G C Majority											
		8110		8120		8130		8140		8150	
1094		C T C C A A G A G C A T T T T A T A A T C A A T G C T G A T G A T T A T C A A A T A G T A A A A G C 2603_all.seq									
1867		C T C C A A G A G C A T T T T A T A A T C A A T G C T G A T G A T T A T C A A A T A G T A A A A G C 18rs21_all.seq									
1099		C T C C A A G A G C A T T T T A T A A T C A A T G C T G A T G A T T A T C A A A T A G T A A A A G C coh1_all.seq									
1099		C T C C A A G A G C A T T T T A T A A T C A A T G C T G A T G A T T A T C A A A T A G T A A A A G C cjb111_all.seq									
1101		C T C C A A G A G C A T T T T A T A A T C A A T G C T G A T G A T T A T C A A A T A G T A A A A G C nem316_all.seq									
1100		C T C C A A G A G C A T T T T A T A A T C A A T G C T G A T G A T T A T C A A A T A G T A A A A G C a909_all.seq									
A G A T G G A G A G A G T T T T A A A C T G T T T T C G G A T A G A A A A G T T C C T G T T A C T G Majority											
		8160		8170		8180		8190		8200	
1144		A G A T G G A G A G A G T T T T A A A C T G T T T T C G G A T A G A A A A G T T C C T G T T A C T G 2603_all.seq									
1917		A G A T G G A G A G A G T T T T A A A C T G T T T T C G G A T A G A A A A G T T C C T G T T A C T G 18rs21_all.seq									
1149		A G A T G G A G A G A G T T T T A A A C T G T T T T C G G A T A G A A A A G T T C C T G T T A C T G coh1_all.seq									
1149		A G A T G G A G A G A G T T T T A A A C T G T T T T C G G A T A G A A A A G T T C C T G T T A C T G cjb111_all.seq									
1151		A G A T G G A G A G A G T T T T A A A C T G T T T T C G G A T A G A A A A G T T C C T G T T A C T G nem316_all.seq									
1150		A G A T G G A G A G A G T T T T A A A C T G T T T T C G G A T A G A A A A G T T C C T G T T A C T G a909_all.seq									
G A G G A A C G A C A C A A G C A G C T T A T C G A G T A C C G C A A A A T C A A C T C T C T G T A Majority											
		8210		8220		8230		8240		8250	
194		G A G G A A C G A C A C A A G C A G C T T A T C G A G T A C C G C A A A A T C A A C T C T C T G T A 2603_all.seq									
967		G A G G A A C G A C A C A A G C A G C T T A T C G A G T A C C G C A A A A T C A A C T C T C T G T A 18rs21_all.seq									
199		G A G G A A C G A C A C A A G C A G C T T A T C G A G T A C C G C A A A A T C A A C T C T C T G T A coh1_all.seq									
199		G A G G A A C G A C A C A A G C A G C T T A T C G A G T A C C G C A A A A T C A A C T C T C T G T A cjb111_all.seq									
201		G A G G A A C G A C A C A A G C A G C T T A T C G A G T A C C G C A A A A T C A A C T C T C T G T A nem316_all.seq									
200		G A G G A A C G A C A C A A G C A G C T T A T C G A G T A C C G C A A A A T C A A C T C T C T G T A a909_all.seq									
A T G A G T A A T G A G G G A T A T G C A A T T A A T A G T G G A T A T A T T T A T C T C T A T T G Majority											
		8260		8270		8280		8290		8300	
244		A T G A G T A A T G A G G G A T A T G C A A T T A A T A G T G G A T A T A T T T A T C T C T A T T G 2603_all.seq									
017		A T G A G T A A T G A G G G A T A T G C A A T T A A T A G T G G A T A T A T T T A T C T C T A T T G 18rs21_all.seq									
249		A T G A G T A A T G A G G G A T A T G C A A T T A A T A G T G G A T A T A T T T A T C T C T A T T G coh1_all.seq									
249		A T G A G T A A T G A G G G A T A T G C A A T T A A T A G T G G A T A T A T T T A T C T C T A T T G cjb111_all.seq									
251		A T G A G T A A T G A G G G A T A T G C A A T T A A T A G T G G A T A T A T T T A T C T C T A T T G nem316_all.seq									
250		A T G A G T A A T G A G G G A T A T G C A A T T A A T A G T G G A T A T A T T T A T C T C T A T T G a909_all.seq									
G A G A G A T T A C A A C T G G G T C T A T C C A T T T G A T C C T A A G A C A A A G A A A G T T T Majority											
		8310		8320		8330		8340		8350	
294		G A G A G A T T A C A A C T G G G T C T A T C C A T T T G A T C C T A A G A C A A A G A A A G T T T 2603_all.seq									
067		G A G A G A T T A C A A C T G G G T C T A T C C A T T T G A T C C T A A G A C A A A G A A A G T T T 18rs21_all.seq									
299		G A G A G A T T A C A A C T G G G T C T A T C C A T T T G A T C C T A A G A C A A A G A A A G T T T coh1_all.seq									
299		G A G A G A T T A C A A C T G G G T C T A T C C A T T T G A T C C T A A G A C A A A G A A A G T T T cjb111_all.seq									
301		G A G A G A T T A C A A C T G G G T C T A T C C A T T T G A T C C T A A G A C A A A G A A A G T T T nem316_all.seq									
300		G A G A G A T T A C A A C T G G G T C T A T C C A T T T G A T C C T A A G A C A A A G A A A G T T T a909_all.seq									
C T G C A A C G A A A G A A A T C A A A A C T C A T G C G T G A G C C A A C A A C A T T A T A C T T T Majority											
		8360		8370		8380		8390		8400	
344		C T G C A A C G A A A C A A A T C A A A A C T C A T G C G T G A G C C A A C A A C A T T A T A C T T T 2603_all.seq									
117		C T G C A A C G A A A C A A A T C A A A A C T C A T G G T G A G C C A A C A A C A T T A T A C T T T 18rs21_all.seq									
349		C T G C A A C G A A A C A A A T C A A A A C T C A T G G T G A G C C A A C A A C A T T A T A C T T T coh1_all.seq									
349		C T G C A A C G A A A C A A A T C A A A A C T C A T G G T G A G C C A A C A A C A T T A T A C T T T cjb111_all.seq									
351		C T G C A A C G A A A C A A A T C A A A A C T C A T G G T G A G C C A A C A A C A T T A T A C T T T nem316_all.seq									
350		C T G C A A C G A A A C A A A T C A A A A C T C A T G G T G A G C C A A C A A C A T T A T A C T T T a909_all.seq									

FIGURE 18 W

42/487

A A T G C A A A T A T A A G A C C T A A A G G T T A T G A C A T T T T T A C T G T T G C G A T T G C Majority	
8410	8420
8394 A A T G C A A A T A T A A G A C C T A A A G G T T A T G A C A T T T T T A C T G T T G C G A T T G C 2603_all.seq	
8167 A A T G C A A A T A T A A G A C C T A A A G G T T A T G A C A T T T T T A C T G T T G C G A T T G C 18rs21_all.seq	
8399 A A T G C A A A T A T A A G A C C T A A A G G T T A T G A C A T T T T T A C T G T T G C G A T T G C coh1_all.seq	
8399 A A T G C A A A T A T A A G A C C T A A A G G T T A T G A C A T T T T T A C T G T T G C G A T T G C cjb111_all.seq	
8401 A A T G C A A A T A T A A G A C C T A A A G G T T A T G A C A T T T T T A C T G T T G C G A T T G C nem316_all.seq	
8400 A A T G C A A A T A T A A G A C C T A A A G G T T A T G A C A T T T T T A C T G T T G C G A T T G C a909_all.seq	
T G T A A A C G G A G A T C C T G G T G C A A C T C C T C T T G A A G C T G A G A A A T T T A T G C Majority	
8460	8470
8444 T G T A A A C G G A G A T C C T G G T G C A A C T C C T C T T G A A G C T G A G A A A T T T A T G C 2603_all.seq	
8217 T G T A A A C G G A G A T C C T G G T G C A A C T C C T C T T G A A G C T G A G A A A T T T A T G C 18rs21_all.seq	
8449 T G T A A A C G G A G A T C C T G G T G C A A C T C C T C T T G A A G C T G A G A A A T T T A T G C coh1_all.seq	
8449 T G T A A A C G G A G A T C C T G G T G C A A C T C C T C T T G A A G C T G A G A A A T T T A T G C cjb111_all.seq	
8451 T G T A A A C G G A G A T C C T G G T G C A A C T C C T C T T G A A G C T G A G A A A T T T A T G C nem316_all.seq	
8450 T G T A A A C G G A G A T C C T G G T G C A A C T C C T C T T G A A G C T G A G A A A T T T A T G C a909_all.seq	
A A T C A A T A T C A A G T A A A A C A G A A A A T T A T A C T A A T G T T G A T G A T A C A A A T Majority	
8510	8520
8494 A A T C A A T A T C A A G T A A A A C A G A A A A T T A T A C T A A T G T T G A T G A T A C A A A T 2603_all.seq	
8267 A A T C A A T A T C A A G T A A A A C A G A A A A T T A T A C T A A T G T T G A T G A T A C A A A T 18rs21_all.seq	
8499 A A T C A A T A T C A A G T A A A A C A G A A A A T T A T A C T A A T G T T G A T G A T A C A A A T coh1_all.seq	
8499 A A T C A A T A T C A A G T A A A A C A G A A A A T T A T A C T A A T G T T G A T G A T A C A A A T cjb111_all.seq	
3501 A A T C A A T A T C A A G T A A A A C A G A A A A T T A T A C T A A T G T T G A T G A T A C A A A T nem316_all.seq	
3500 A A T C A A T A T C A A G T A A A A C A G A A A A T T A T A C T A A T G T T G A T G A T A C A A A T a909_all.seq	
A A A A T T T A T G A T G A C C T A A A T A A A T A C T T T A A A A C A A T T G T T G A G C A A A A Majority	
8560	8570
3544 A A A A T T T A T G A T G A C C T A A A T A A A T A C T T T A A A A C A A T T G T T G A G C A A A A 2603_all.seq	
3317 A A A A T T T A T G A T G A C C T A A A T A A A T A C T T T A A A A C A A T T G T T G A G C A A A A 18rs21_all.seq	
3549 A A A A T T T A T G A T G A C C T A A A T A A A T A C T T T A A A A C A A T T G T T G A G C A A A A coh1_all.seq	
3549 A A A A T T T A T G A T G A C C T A A A T A A A T A C T T T A A A A C A A T T G T T G A G C A A A A cjb111_all.seq	
3551 A A A A T T T A T G A T G A C C T A A A T A A A T A C T T T A A A A C A A T T G T T G A G C A A A A nem316_all.seq	
3550 A A A A T T T A T G A T G A C C T A A A T A A A T A C T T T A A A A C A A T T G T T G A G C A A A A a909_all.seq	
A C A T T C T A T T C T T G A T G C A A A T G T G A C T G A T C C T A T G C G A G A G A T G A T T C Majority	
8610	8620
1594 A C A T T C T A T T C T T G A T G C A A A T G T G A C T G A T C C T A T G C G A G A G A T G A T T C 2603_all.seq	
1367 A C A T T C T A T T C T T G A T G C A A A T G T G A C T G A T C C T A T G C G A G A G A T G A T T C 18rs21_all.seq	
1599 A C A T T C T A T T C T T G A T G C A A A T G T G A C T G A T C C T A T G C G A G A G A T G A T T C coh1_all.seq	
1599 A C A T T C T A T T C T T G A T G C A A A T G T G A C T G A T C C T A T G C G A G A G A T G A T T C cjb111_all.seq	
1601 A C A T T C T A T T C T T G A T G C A A A T G T G A C T G A T C C T A T G C G A G A G A T G A T T C nem316_all.seq	
1600 A C A T T C T A T T C T T G A T G C A A A T G T G A C T G A T C C T A T G C G A G A G A T G A T T C a909_all.seq	
A A T T C C A A T T A A A A A A T G G T C A A A G T T T T A C A C A T G A T G A T T A C G T T T T C Majority	
8660	8670
1644 A A T T C C A A T T A A A A A A T G G T C A A A G T T T T A C A C A T G A T G A T T A C G T T T T C 2603_all.seq	
1417 A A T T C C A A T T A A A A A A T G G T C A A A G T T T T A C A C A T G A T G A T T A C G T T T T C 18rs21_all.seq	
1649 A A T T C C A A T T A A A A A A T G G T C A A A G T T T T A C A C A T G A T G A T T A C G T T T T C coh1_all.seq	
1649 A A T T C C A A T T A A A A A A T G G T C A A A G T T T T A C A C A T G A T G A T T A C G T T T T C cjb111_all.seq	
1651 A A T T C C A A T T A A A A A A T G G T C A A A G T T T T A C A C A T G A T G A T T A C G T T T T C nem316_all.seq	
1650 A A T T C C A A T T A A A A A A T G G T C A A A G T T T T A C A C A T G A T G A T T A C G T T T T C a909_all.seq	
G T T G C A A A T G A T G C C A C T C A A T T A A A A A A T G G T G T G C C T C T T G C T G C A C C Majority	
8710	8720
1694 G T T G C A A A T G A T G C C A C T C A A T T A A A A A A T G G T G T G C C T C T T G C T G C A C C 2603_all.seq	
1467 G T T G C A A A T G A T G C C A C T C A A T T A A A A A A T G G T G T G C C T C T T G C T G C A C C 18rs21_all.seq	
1699 G T T G C A A A T G A T G C C A C T C A A T T A A A A A A T G G T G T G C C T C T T G C T G C A C C coh1_all.seq	
1699 G T T G C A A A T G A T G C C A C T C A A T T A A A A A A T G G T G T G C C T C T T G C T G C A C C cjb111_all.seq	
701 G T T G C A A A T G A T G C C A C T C A A T T A A A A A A T G G T G T G C C T C T T G C T G C A C C nem316_all.seq	
700 G T T G C A A A T G A T G C C A C T C A A T T A A A A A A T G G T G T G C C T C T T G C T G C A C C a909_all.seq	

FIGURE 18 X

A A A C A G T G A T G C G C G A A T T T T A A A A G A T G T T A C A G T G A C T T A T C A T A A G A Majority										
8760		8770		8780		8790		8800		
8744	A A A C A G T G A T G C G C G A A T T T T A A A A G A T G T T A C A G T G A C T T A T C A T A A G A									2603_all.seq
8517	A A A C A G T G A T G C G C G A A T T T T A A A A G A T G T T A C A G T G A C T T A T C A T A A G A									18rs21_all.seq
8749	A A A C A G T G A T G C G C G A A T T T T A A A A G A T G T T A C A G T G A C T T A T C A T A A G A									cohl_all.seq
8749	A A A C A G T G A T G C G C G A A T T T T A A A A G A T G T T A C A G T G A C T T A T C A T A A G A									cjb111_all.seq
8751	A A A C A G T G A T G C G C G A A T T T T A A A A G A T G T T A C A G T G A C T T A T C A T A A G A									nem316_all.seq
8750	A A A C A G T G A T G C G C G A A T T T T A A A A G A T G T T A C A G T G A C T T A T C A T A A G A									a909_all.seq
C A T C T C A A A C C A T C A A A A T C A A T C A T T T G A A C T T A G G A A G T G C A C A A A A A Majority										
8810		8820		8830		8840		8850		
8794	C A T C T C A A A C C A T C A A A A T C A A T C A T T T G A A C T T A G G A A G T G C A C A A A A A									2603_all.seq
8567	C A T C T C A A A C C A T C A A A A T C A A T C A T T T G A A C T T A G G A A G T G C A C A A A A A									18rs21_all.seq
8799	C A T C T C A A A C C A T C A A A A T C A A T C A T T T G A A C T T A G G A A G T G C A C A A A A A									cohl_all.seq
8799	C A T C T C A A A C C A T C A A A A T C A A T C A T T T G A A C T T A G G A A G T G C A C A A A A A									cjb111_all.seq
8801	C A T C T C A A A C C A T C A A A A T C A A T C A T T T G A A C T T A G G A A G T G C A C A A A A A									nem316_all.seq
8800	C A T C T C A A A C C A T C A A A A T C A A T C A T T T G A A C T T A G G A A G T G C A C A A A A A									a909_all.seq
G T A G T T C T T A C C T A T G A T G T A C G T T T A A A A G A T A A C T A T A T A A G T A A C A A Majority										
8860		8870		8880		8890		8900		
8844	G T A G T T C T T A C C T A T G A T G T A C G T T T A A A A G A T A A C T A T A T A A G T A A C A A									2603_all.seq
8617	G T A G T T C T T A C C T A T G A T G T A C G T T T A A A A G A T A A C T A T A T A A G T A A C A A									18rs21_all.seq
8849	G T A G T T C T T A C C T A T G A T G T A C G T T T A A A A G A T A A C T A T A T A A G T A A C A A									cohl_all.seq
8849	G T A G T T C T T A C C T A T G A T G T A C G T T T A A A A G A T A A C T A T A T A A G T A A C A A									cjb111_all.seq
8851	G T A G T T C T T A C C T A T G A T G T A C G T T T A A A A G A T A A C T A T A T A A G T A A C A A									nem316_all.seq
8850	G T A G T T C T T A C C T A T G A T G T A C G T T T A A A A G A T A A C T A T A T A A G T A A C A A									a909_all.seq
A T T T T A C A A T A C A A A T A A T C G T A C A A C G C T A A G T C C G A A G A G T G A A A A A G Majority										
8910		8920		8930		8940		8950		
8894	A T T T T A C A A T A C A A A T A A T C G T A C A A C G C T A A G T C C G A A G A G T G A A A A A G									2603_all.seq
8667	A T T T T A C A A T A C A A A T A A T C G T A C A A C G C T A A G T C C G A A G A G T G A A A A A G									18rs21_all.seq
8899	A T T T T A C A A T A C A A A T A A T C G T A C A A C G C T A A G T C C G A A G A G T G A A A A A G									cohl_all.seq
8899	A T T T T A C A A T A C A A A T A A T C G T A C A A C G C T A A G T C C G A A G A G T G A A A A A G									cjb111_all.seq
8901	A T T T T A C A A T A C A A A T A A T C G T A C A A C G C T A A G T C C G A A G A G T G A A A A A G									nem316_all.seq
8900	A T T T T A C A A T A C A A A T A A T C G T A C A A C G C T A A G T C C G A A G A G T G A A A A A G									a909_all.seq
A A C C A A A T A C T A T T C G T G A T T T C C C A A T T C C C A A A A T T C G T G A T G T T C G T Majority										
8960		8970		8980		8990		9000		
3944	A A C C A A A T A C T A T T C G T G A T T T C C C A A T T C C C A A A A T T C G T G A T G T T C G T									2603_all.seq
3717	A A C C A A A T A C T A T T C G T G A T T T C C C A A T T C C C A A A A T T C G T G A T G T T C G T									18rs21_all.seq
3949	A A C C A A A T A C T A T T C G T G A T T T C C C A A T T C C C A A A A T T C G T G A T G T T C G T									cohl_all.seq
3949	A A C C A A A T A C T A T T C G T G A T T T C C C A A T T C C C A A A A T T C G T G A T G T T C G T									cjb111_all.seq
3951	A A C C A A A T A C T A T T C G T G A T T T C C C A A T T C C C A A A A T T C G T G A T G T T C G T									nem316_all.seq
3950	A A C C A A A T A C T A T T C G T G A T T T C C C A A T T C C C A A A A T T C G T G A T G T T C G T									a909_all.seq
G A G T T T C C G G T A C T A A C C A T C A G T A A T C A G A A C A A A A T G G G T G A G G T T G A Majority										
9010		9020		9030		9040		9050		
3994	G A G T T T C C G G T A C T A A C C A T C A G T A A T C A G A A C A A A A T G G G T G A G G T T G A									2603_all.seq
3767	G A G T T T C C G G T A C T A A C C A T C A G T A A T C A G A A C A A A A T G G G T G A G G T T G A									18rs21_all.seq
3999	G A G T T T C C G G T A C T A A C C A T C A G T A A T C A G A A C A A A A T G G G T G A G G T T G A									cohl_all.seq
3999	G A G T T T C C G G T A C T A A C C A T C A G T A A T C A G A A C A A A A T G G G T G A G G T T G A									cjb111_all.seq
3001	G A G T T T C C G G T A C T A A C C A T C A G T A A T C A G A A C A A A A T G G G T G A G G T T G A									nem316_all.seq
3000	G A G T T T C C G G T A C T A A C C A T C A G T A A T C A G A A C A A A A T G G G T G A G G T T G A									a909_all.seq
A T T T A T T A A A G T T A A T A A A G A C A A A C A T T C A G A A T C G C T T T T G C G A C C T A Majority										
9060		9070		9080		9090		9100		
3044	A T T T A T T A A A G T T A A T A A A G A C A A A C A T T C A G A A T C G C T T T T G C G A C C T A									2603_all.seq
3817	A T T T A T T A A A G T T A A T A A A G A C A A A C A T T C A G A A T C G C T T T T G C G A C C T A									18rs21_all.seq
3049	A T T T A T T A A A G T T A A T A A A G A C A A A C A T T C A G A A T C G C T T T T G C G A C C T A									cohl_all.seq
3049	A T T T A T T A A A G T T A A T A A A G A C A A A C A T T C A G A A T C G C T T T T G C G A C C T A									cjb111_all.seq
3051	A T T T A T T A A A G T T A A T A A A G A C A A A C A T T C A G A A T C G C T T T T G C G A C C T A									nem316_all.seq
3050	A T T T A T T A A A G T T A A T A A A G A C A A A C A T T C A G A A T C G C T T T T G C G A C C T A									a909_all.seq

FIGURE 18 Y

44/487

A G T T T C A A C T T T C A G A T T A C A A A A A A G A T T T T C T G G G T A T A A G C A A T T T G T T Majority											
		9110		9120		9130		9140		9150	
9094	A	G	T	T	T	C	A	A	C	T	T
8867	A	G	T	T	T	C	A	A	C	T	T
9099	A	G	T	T	T	C	A	A	C	T	T
9099	A	G	T	T	T	C	A	A	C	T	T
9101	A	G	T	T	T	C	A	A	C	T	T
9100	A	G	T	T	T	C	A	A	C	T	T
C C A G A G G G A A G T G A T G T T A C A A C A A A G A A T G A T G G T A A A A T T T A T T T T A A Majority											
		9160		9170		9180		9190		9200	
9144	C	C	A	G	A	G	G	A	A	G	T
8917	C	C	A	G	A	G	G	A	A	G	T
9149	C	C	A	G	A	G	G	A	A	G	T
9149	C	C	A	G	A	G	G	A	A	G	T
9151	C	C	A	G	A	G	G	A	A	G	T
9150	C	C	A	G	A	G	G	A	A	G	T
A G C A C T T C A A G A T G G T A A C T A T A A A T T A T A T G A A A T T T C A A G T C C A G A T C Majority											
		9210		9220		9230		9240		9250	
9194	A	G	C	A	C	T	T	C	A	A	G
9367	A	G	C	A	C	T	T	C	A	A	G
9199	A	G	C	A	C	T	T	C	A	A	G
9199	A	G	C	A	C	T	T	C	A	A	G
9201	A	G	C	A	C	T	T	C	A	A	G
9200	A	G	C	A	C	T	T	C	A	A	G
G C T A T A T A G A G G T T A A A A C G A A A C C T G T T G T G A C A T T T A C A A T T C A A A A T Majority											
		9260		9270		9280		9290		9300	
9244	G	C	T	A	T	A	T	A	G	A	G
9017	G	C	T	A	T	A	T	A	G	A	G
9249	G	C	T	A	T	A	T	A	G	A	G
9249	G	C	T	A	T	A	T	A	G	A	G
9251	G	C	T	A	T	A	T	A	G	A	G
9250	G	C	T	A	T	A	T	A	G	A	G
G G A G A A G T T A C G A A C C T G A A A G C A G A T C C A A A T G C T A A T A A A A A T C A A A T Majority											
		9310		9320		9330		9340		9350	
294	G	G	A	G	A	A	G	T	T	A	C
067	G	G	A	G	A	A	G	T	T	A	C
299	G	G	A	G	A	A	G	T	T	A	C
299	G	G	A	G	A	A	G	T	T	A	C
301	G	G	A	G	A	A	G	T	T	A	C
300	G	G	A	G	A	A	G	T	T	A	C
C G G G T A T C T T G A A G G A A A T G G T A A A C A T C T T A T T A C C A A C A C T C C C A A A C Majority											
		9360		9370		9380		9390		9400	
344	C	G	G	G	T	A	T	C	T	T	G
117	C	G	G	G	T	A	T	C	T	T	G
349	C	G	G	G	T	A	T	C	T	T	G
349	C	G	G	G	T	A	T	C	T	T	G
351	C	G	G	G	T	A	T	C	T	T	G
350	C	G	G	G	T	A	T	C	T	T	G
G C C C A C C A G G T G T T T T T C C T A A A A C A G G G G G A A T T G G T A C A A T T G T C T A T Majority											
		9410		9420		9430		9440		9450	
394	G	C	C	C	A	C	C	A	G	G	T
167	G	C	C	C	A	C	C	A	G	G	T
399	G	C	C	C	A	C	C	A	G	G	T
399	G	C	C	C	A	C	C	A	G	G	T
401	G	C	C	C	A	C	C	A	G	G	T
400	G	C	C	C	A	C	C	A	G	G	T

FIGURE 18 Z

45/487

A T A T T A G C T T G C T T C T A C T T T T A T G A T A C T T A C C A T T T G T T C T T T C C G T C G Majority										
9460		9470		9480		9490		9500		
9444	A T A T T A G C T T G C T T C T A C T T T T A T G A T A C T T A C C A T T T G T T C T T T C C G T C G									2603_all.seq
9217	A T A T T A G C T T G C T T C T A C T T T T A T G A T A C T T A C C A T T T G T T C T T T C C G T C G									18rs21_all.seq
9449	A T A T T A G C T T G C T T C T A C T T T T A T G A T A C T T A C C A T T T G T T C T T T C C G T C G									cohl_all.seq
9449	A T A T T A G C T T G C T T C T A C T T T T A T G A T A C T T A C C A T T T G T T C T T T C C G T C G									cjb111_all.seq
9451	A T A T T A G C T T G C T T C T A C T T T T A T G A T A C T T A C C A T T T G T T C T T T C C G T C G									nem316_all.seq
9450	A T A T T A G C T T G C T T C T A C T T T T A T G A T A C T T A C C A T T T G T T C T T T C C G T C G									a909_all.seq
T A A A C A A T T G T A A G C T G T C G T T G A A A T T A T T A A A T A A T A G A A A A T G A C T A Majority										
9510		9520		9530		9540		9550		
9494	T A A A C A A T T G T A A G C T G T C G T T G A A A T T A T T A A A T A A T A G A A A A T G A C T A									2603_all.seq
9267	T A A A C A A T T G T A A G C T G T C G T T G A A A T T A T T A A A T A A T A G A A A A T G A C T A									18rs21_all.seq
9499	T A A A C A A T T G T A A G C T G T C G T T G A A A T T A T T A A A T A A T A G A A A A T G A C T A									cohl_all.seq
9499	T A A A C A A T T G T A A G C T G T C G T T G A A A T T A T T A A A T A A T A G A A A A T G A C T A									cjb111_all.seq
9501	T A A A C A A T T G T A A G C T G T C G T T G A A A T T A T T A A A T A A T A G A A A A T G A C T A									nem316_all.seq
9500	T A A A C A A T T G T A A G C T G T C G T T G A A A T T A T T A A A T A A T A G A A A A T G A C T A									a909_all.seq
G T T T T G C C C T T T C C C T A T T G T C A G T C A G A T T A G T T A T T A T C A A G C T T C T C Majority										
9560		9570		9580		9590		9600		
9544	G T T T T G C C C T T T C C C T A T T G T C A G T C A G A T T A G T T A T T A T C A A G C T T C T C									2603_all.seq
9317	G T T T T G C C C T T T C C C T A T T G T C A G T C A G A T T A G T T A T T A T C A A G C T T C T C									18rs21_all.seq
9549	G T T T T G C C C T T T C C C T A T T G T C A G T C A G A T T A G T T A T T A T C A A G C T T C T C									cohl_all.seq
9549	G T T T T G C C C T T T C C C T A T T G T C A G T C A G A T T A G T T A T T A T C A A G C T T C T C									cjb111_all.seq
9551	G T T T T G C C C T T T C C C T A T T G T C A G T C A G A T T A G T T A T T A T C A A G C T T C T C									nem316_all.seq
9550	G T T T T G C C C T T T C C C T A T T G T C A G T C A G A T T A G T T A T T A T C A A G C T T C T C									a909_all.seq
A T G C C A A T A T T A A T G C C T T T A A A A G A A C A G T T A C A A C C A T T G A C C G T A C G Majority										
9610		9620		9630		9640		9650		
594	A T G C C A A T A T T A A T G C C T T T A A A A G A A C A G T T A C A A C C A T T G A C C G T A C G									2603_all.seq
967	A T G C C A A T A T T A A T G C C T T T A A A A G A A C A G T T A C A A C C A T T G A C C G T A C G									18rs21_all.seq
599	A T G C C A A T A T T A A T G C C T T T A A A A G A A C A G T T A C A A C C A T T G A C C G T A C G									cohl_all.seq
599	A T G C C A A T A T T A A T G C C T T T A A A A G A A C A G T T A C A A C C A T T G A C C G T A C G									cjb111_all.seq
601	A T G C C A A T A T T A A T G C C T T T A A A A G A A C A G T T A C A A C C A T T G A C C G T A C G									nem316_all.seq
600	A T G C C A A T A T T A A T G C C T T T A A A A G A A C A G T T A C A A C C A T T G A C C G T A C G									a909_all.seq
G A G A T T C A A C G A C G C T T A G C T C T G G C T A A A G C C T A C A A T G C T A G T A T T T C Majority										
9660		9670		9680		9690		9700		
644	G A G A T T C A A C G A C G G T T A G C T C T G G C T A A A G C C T A C A A T G C T A G T A T T T C									2603_all.seq
417	G A G A T T C A A C G A C G G T T A G C T C T G G C T A A A G C C T A C A A T G C T A G T A T T T C									18rs21_all.seq
649	G A G A T T C A A C G A C G G T T A G C T C T G G C T A A A G C C T A C A A T G C T A G T A T T T C									cohl_all.seq
649	G A G A T T C A A C G A C G G T T A G C T C T G G C T A A A G C C T A C A A T G C T A G T A T T T C									cjb111_all.seq
651	G A G A T T C A A C G A C G G T T A G C T C T G G C T A A A G C C T A C A A T G C T A G T A T T T C									nem316_all.seq
650	G A G A T T C A A C G A C G G T T A G C T C T G G C T A A A G C C T A C A A T G C T A G T A T T T C									a909_all.seq
T G G A A C A A G T A G T C A G T C G A C T C A A T C T G T G C T G A G A G A T T C T T A T T C T G Majority										
9710		9720		9730		9740		9750		
394	T G G A A C A A G T A G T C A G T C G A C T C A A T C T G T G C T G A G A G A T T C T T A T T C T G									2603_all.seq
467	T G G A A C A A G T A G T C A G T C G A C T C A A T C T G T G C T G A G A G A T T C T T A T T C T G									18rs21_all.seq
399	T G G A A C A A G T A G T C A G T C G A C T C A A T C T G T G C T G A G A G A T T C T T A T T C T G									cohl_all.seq
399	T G G A A C A A G T A G T C A G T C G A C T C A A T C T G T G C T G A G A G A T T C T T A T T C T G									cjb111_all.seq
701	T G G A A C A A G T A G T C A G T C G A C T C A A T C T G T G C T G A G A G A T T C T T A T T C T G									nem316_all.seq
700	T G G A A C A A G T A G T C A G T C G A C T C A A T C T G T G C T G A G A G A T T C T T A T T C T G									a909_all.seq
A G G A G C A A A A A A G C G A C G G T G G A C T G A A T A C G C T A G C A T G T T A G A A G T C A Majority										
9760		9770		9780		9790		9800		
144	A G G A G C A A A A A A G C G A C G G T G G A C T G A A T A C G C T A G C A T G T T A G A A G T C A									2603_all.seq
517	A G G A G C A A A A A A G C G A C G G T G G A C T G A A T A C G C T A G C A T G T T A G A A G T C A									18rs21_all.seq
149	A G G A G C A A A A A A G C G A C G G T G G A C T G A A T A C G C T A G C A T G T T A G A A G T C A									cohl_all.seq
149	A G G A G C A A A A A A G C G A C G G T G G A C T G A A T A C G C T A G C A T G T T A G A A G T C A									cjb111_all.seq
151	A G G A G C A A A A A A G C G A C G G T G G A C T G A A T A C G C T A G C A T G T T A G A A G T C A									nem316_all.seq
150	A G G A G C A A A A A A G C G A C G G T G G A C T G A A T A C G C T A G C A T G T T A G A A G T C A									a909_all.seq

FIGURE 18 AA

46/487

WO 2006/078318

Alignment Report of AI-1 alignment, using J. Hein method with Weighted residue weight table.
Thursday, July 29, 2004 5:46 PM

PCT/US2005/027239

Page 29

G A G A G C A G G T T G A C C A T G T G A T G A T T C C A A A A A T C A A T C A G G A T T T A C C A Majority									
9810		9820		9830		9840		9850	
1794	G	A	G	A	G	C	A	G	G
1567	G	A	G	A	G	C	A	G	G
1799	G	A	G	A	G	C	A	G	G
1799	G	A	G	A	G	C	A	G	G
1801	G	A	G	A	G	C	A	G	G
1800	G	A	G	A	G	C	A	G	G
A T C T A C G C T G G T T C A G A A G A G G A C A A T C T G C A A C G G G G A G T T G G T C A T C T Majority									
9860		9870		9880		9890		9900	
1844	A	T	C	T	A	C	G	C	T
1617	A	T	C	T	A	C	G	C	T
1849	A	T	C	T	A	C	G	C	T
1849	A	T	C	T	A	C	G	C	T
1851	A	T	C	T	A	C	G	C	T
1850	A	T	C	T	A	C	G	C	T
A C A A C G G A T A A G T T T G C C G A T T G G A G G G G C T T C T A C A C A T G C G G T C T T G A Majority									
9910		9920		9930		9940		9950	
1894	A	C	A	A	C	G	G	A	T
1667	A	C	A	A	C	G	G	A	T
1899	A	C	A	A	C	G	G	A	T
1899	A	C	A	A	C	G	G	A	T
901	A	C	A	A	C	G	G	A	T
900	A	C	A	A	C	G	G	A	T
G C C G T C A A A G A G G T A T G C C A G C T G C T C G G T T G T T T G C G G A T T T G G A T A A G Majority									
9960		9970		9980		9990		10000	
944	G	C	C	G	T	C	A	A	A
717	G	C	C	G	T	C	A	A	A
949	G	C	C	G	T	C	A	A	A
949	G	C	C	G	T	C	A	A	A
951	G	C	C	G	T	C	A	A	A
950	G	C	C	G	T	C	A	A	A
A T G A A A A A A G G T G A T T A T T T T A T G T T A C C A A T C T G A A A G A A A C C T T G G C Majority									
10010		10020		10030		10040		10050	
994	A	T	G	A	A	A	A	A	G
767	A	T	G	A	A	A	A	A	G
999	A	T	G	A	A	A	A	A	G
999	A	T	G	A	A	A	A	A	G
0001	A	T	G	A	A	A	A	A	G
0000	A	T	G	A	A	A	A	A	G
T T A T C A A G T G G A T C G T A T C A T G G T G A T T G A A C C T A G C C A A T T G C A T G C C C Majority									
10060		10070		10080		10090		10100	
0044	T	T	A	T	C	A	A	G	T
817	T	T	A	T	C	A	A	G	T
0049	T	T	A	T	C	A	A	G	T
0049	T	T	A	T	C	A	A	G	T
0051	T	T	A	T	C	A	A	G	T
0050	T	T	A	T	C	A	A	G	T
T C A G C A T T G A A G A G G A T A A A G A T T A T G T T A C C C T T C T G A C C T G T A C A C C T Majority									
10110		10120		10130		10140		10150	
0094	T	C	A	G	C	A	T	T	G
867	T	C	A	G	C	A	T	T	G
0099	T	C	A	G	C	A	T	T	G
0099	T	C	A	G	C	A	T	T	G
0101	T	C	A	G	C	A	T	T	G
0100	T	C	A	G	C	A	T	T	G

FIGURE 18 AB

47/487

T A T A T G G C T C T T T G T C A A C T G T A A T G G G T G A C T T A T C A T T A A C A A C G A G		Majority
10160	10170	10180
10144	T A T A T G G G C T C T T T G T C A A C T G T A A T G G G T G A C T T A T C A T T A A C A A C G A G	2603_all.seq
9917	T A T A T G G G C T C T T T G T C A A C T G T A A T G G G T G A C T T A T C A T T A A C A A C G A G	18rs21_all.seq
10149	T A T A T G G G C T C T T T G T C A A C T G T A A T G G G T G A C T T A T C A T T A A C A A C G A G	cohl_all.seq
10149	T A T A T G G G C T C T T T G T C A A C T G T A A T G G G T G A C T T A T C A T T A A C A A C G A G	cjb111_all.seq
10151	T A T A T G G G C T C T T T G T C A A C T G T A A T G G G T G A C T T A T C A T T A A C A A C G A G	nen316_all.seq
10150	T A T A T G G G C T C T T T G T C A A C T G T A A T G G G T G A C T T A T C A T T A A C A A C G A G	a909_all.seq
A G A G A A T C A G C T T G G T T C T C T C T T T T T G G A T G T T C A A A G C G A T G A G A A		
10210	10220	10230
10194	A G A G A A T C A G C T T G G T T C T C T C T T T T T G G A T G T T C A A A G C G A T G A G A A	2603_all.seq
9957	A G A G A A T C A G C T T G G T T C T C T C T C T T T T T G G A T G T T C A A A G C G A T G A G A A	18rs21_all.seq
10199	A G A G A A T C A G C T T G G T T C T C T C T C T T T T T G G A T G T T C A A A G C G A T G A G A A	cohl_all.seq
10199	A G A G A A T C A G C T T G G T T C T C T C T C T T T T T G G A T G T T C A A A G C G A T G A G A A	cjb111_all.seq
10201	A G A G A A T C A G C T T G G T T C T C T C T C T T T T T G G A T G T T C A A A G C G A T G A G A A	nen316_all.seq
10200	A G A G A A T C A G C T T G G T T C T C T C T C T T T T T G G A T G T T C A A A G C G A T G A G A A	a909_all.seq
T T T T G C T T T T A A A A T T T T T A A A G T T G C G A A A G C C A A A G C C T T G T C G C T T G		
10260	10270	10280
10244	T T T T G C T T T T A A A A T T T T T A A A G T T G C G A A A G C C A A A G C C T T G T C G C T T G	2603_all.seq
10017	T T T T G C T T T T A A A A T T T T T A A A G T T G C G A A A G C C A A A G C C T T G T C G C T T G	18rs21_all.seq
10249	T T T T G C T T T T A A A A T T T T T A A A G T T G C G A A A G C C A A A G C C T T G T C G C T T G	cohl_all.seq
10249	T T T T G C T T T T A A A A T T T T T A A A G T T G C G A A A G C C A A A G C C T T G T C G C T T G	cjb111_all.seq
10251	T T T T G C T T T T A A A A T T T T T A A A G T T G C G A A A G C C A A A G C C T T G T C G C T T G	nen316_all.seq
10250	T T T T G C T T T T A A A A T T T T T A A A G T T G C G A A A G C C A A A G C C T T G T C G C T T G	a909_all.seq
A T G T C T T T G A T A A G C T T G T T G C T T C T A A T T T G G C A T T G G A A A G G G A		
10310	10320	10330
10294	A T G T C T T T G A T A A G C T T G T T G C T T C T A A T T T G G C A T T G G A A A G G G A	2603_all.seq
10067	A T G T C T T T G A T A A G C T T G T T G C T T C T A A T T T G G C A T T G G A A A G G G A	18rs21_all.seq
10299	A T G T C T T T G A T A A G C T T G T T G C T T C T A A T T T G G C A T T G G A A A G G G A	cohl_all.seq
10299	A T G T C T T T G A T A A G C T T G T T G C T T C T A A T T T G G C A T T G G A A A G G G A	cjb111_all.seq
10301	A T G T C T T T G A T A A G C T T G T T G C T T C T A A T T T G G C A T T G G A A A G G G A	nen316_all.seq
10300	A T G T C T T T G A T A A G C T T G T T G C T T C T A A T T T G G C A T T G G A A A G G G A	a909_all.seq
T A G C T G A G C C C T T T G T G A T G T A A G T T T T G T A C T T C A T G A A G G T C C T A A A		
10360	10370	10380
10344	T A G C T G A G C C C T T T G T G A T G T A A G T T T T G T A C T T C A T G A A G G T C C T A A A	2603_all.seq
10117	T A G C T G A G C C C T T T G T G A T G T A A G T T T T G T A C T T C A T G A A G G T C C T A A A	18rs21_all.seq
10349	T A G C T G A G C C C T T T G T G A T G T A A G T T T T G T A C T T C A T G A A G G T C C T A A A	cohl_all.seq
10349	T A G C T G A G C C C T T T G T G A T G T A A G T T T T G T A C T T C A T G A A G G T C C T A A A	cjb111_all.seq
10351	T A G C T G A G C C C T T T G T G A T G T A A G T T T T G T A C T T C A T G A A G G T C C T A A A	nen316_all.seq
10350	T A G C T G A G C C C T T T G T G A T G T A A G T T T T G T A C T T C A T G A A G G T C C T A A A	a909_all.seq
G A C C A T T T T A A A A G C A G A A T T G C C A G T T G T T A C G T T A T G A T C T A T G A A G C		
10410	10420	10430
0394	G A C C A T T T T A A A A G C A G A A T T G C C A G T T G T T A C G T T A T G A T C T A T G A A G C	2603_all.seq
0167	G A C C A T T T T A A A A G C A G A A T T G C C A G T T G T T A C G T T A T G A T C T A T G A A G C	18rs21_all.seq
0399	G A C C A T T T T A A A A G C A G A A T T G C C A G T T G T T A C G T T A T G A T C T A T G A A G C	cohl_all.seq
0399	G A C C A T T T T A A A A G C A G A A T T G C C A G T T G T T A C G T T A T G A T C T A T G A A G C	cjb111_all.seq
0401	G A C C A T T T T A A A A G C A G A A T T G C C A G T T G T T A C G T T A T G A T C T A T G A A G C	nen316_all.seq
0400	G A C C A T T T T A A A A G C A G A A T T G C C A G T T G T T A C G T T A T G A T C T A T G A A G C	a909_all.seq
T A A A G A A T T C A G T T G T T C C C T T T T C T T A G A A A T G A A A G A G T A T G A G C T G A		
10460	10470	10480
0444	T A A A G A A T T C A G T T G T T C C C T T T T C T T A G A A A T G A A A G A G T A T G A G C T G A	2603_all.seq
0217	T A A A G A A T T C A G T T G T T C C C T T T T C T T A G A A A T G A A A G A G T A T G A G C T G A	18rs21_all.seq
0449	T A A A G A A T T C A G T T G T T C C C T T T T C T T A G A A A T G A A A G A G T A T G A G C T G A	cohl_all.seq
0449	T A A A G A A T T C A G T T G T T C C C T T T T C T T A G A A A T G A A A G A G T A T G A G C T G A	cjb111_all.seq
0451	T A A A G A A T T C A G T T G T T C C C T T T T C T T A G A A A T G A A A G A G T A T G A G C T G A	nen316_all.seq
0450	T A A A G A A T T C A G T T G T T C C C T T T T C T T A G A A A T G A A A G A G T A T G A G C T G A	a909_all.seq

FIGURE 18 AC

TAAAGGTCGTTATAATAGCGGAGCTCATCTGAGAAAACCTTTTATACCTCA Majority									
10510		10520		10530		10540		10550	
10494	TAAAGGTCGTTATAATAGCGGAGCTCATCTGAGAAAACCTTTTATACCTCA	2603_all.seq							
10267	TAAAGGTCGTTATAATAGCGGAGCTCATCTGAGAAAACCTTTTATACCTCA	18rs21_all.seq							
10499	TAAAGGTCGTTATAATAGCGGAGCTCATCTGAGAAAACCTTTTATACCTCA	cohl_all.seq							
10499	TAAAGGTCGTTATAATAGCGGAGCTCATCTGAGAAAACCTTTTATACCTCA	cjb111_all.seq							
10501	TAAAGGTCGTTATAATAGCGGAGCTCATCTGAGAAAACCTTTTATACCTCA	nem316_all.seq							
10500	TAAAGGTCGTTATAATAGCGGAGCTCATCTGAGAAAACCTTTTATACCTCA	a909_all.seq							
AAGTCAGTCTAGCTTTGATATCATGAAGCCATTAGGAGTTATTCCCTTATC Majority									
10560		10570		10580		10590		10600	
10544	AAGTCAGTCTAGCTTTGATATCATGAAGCCATTAGGAGTTATTCCCTTATC	2603_all.seq							
10317	AAGTCAGTCTAGCTTTGATATCATGAAGCCATTAGGAGTTATTCCCTTATC	18rs21_all.seq							
10549	AAGTCAGTCTAGCTTTGATATCATGAAGCCATTAGGAGTTATTCCCTTATC	cohl_all.seq							
10549	AAGTCAGTCTAGCTTTGATATCATGAAGCCATTAGGAGTTATTCCCTTATC	cjb111_all.seq							
10551	AAGTCAGTCTAGCTTTGATATCATGAAGCCATTAGGAGTTATTCCCTTATC	nem316_all.seq							
10550	AAGTCAGTCTAGCTTTGATATCATGAAGCCATTAGGAGTTATTCCCTTATC	a909_all.seq							
TTTTAGTGGCGCGCGATCCATATAGTGATAGATCGAGATATTTAGATCCA Majority									
10610		10620		10630		10640		10650	
10594	TTTTAGTGGCGCGCGATCCATATAGTGATAGATCGAGATATTTAGATCCA	2603_all.seq							
10367	TTTTAGTGGCGCGCGATCCATATAGTGATAGATCGAGATATTTAGATCCA	18rs21_all.seq							
10599	TTTTAGTGGCGCGCGATCCATATAGTGATAGATCGAGATATTTAGATCCA	cohl_all.seq							
10599	TTTTAGTGGCGCGCGATCCATATAGTGATAGATCGAGATATTTAGATCCA	cjb111_all.seq							
10601	TTTTAGTGGCGCGCGATCCATATAGTGATAGATCGAGATATTTAGATCCA	nem316_all.seq							
10600	TTTTAGTGGCGCGCGATCCATATAGTGATAGATCGAGATATTTAGATCCA	a909_all.seq							
AAAGTTCTATCATCCTCTTTTGGCGGCTTTTTCAGCAGATAAATATTA Majority									
10660		10670		10680		10690		10700	
10644	AAAGTTCTATCATCCTCTTTTGGCGGCTTTTTCAGCAGATAAATATTA	2603_all.seq							
10417	AAAGTTCTATCATCCTCTTTTGGCGGCTTTTTCAGCAGATAAATATTA	18rs21_all.seq							
10649	AAAGTTCTATCATCCTCTTTTGGCGGCTTTTTCAGCAGATAAATATTA	cohl_all.seq							
10649	AAAGTTCTATCATCCTCTTTTGGCGGCTTTTTCAGCAGATAAATATTA	cjb111_all.seq							
10651	AAAGTTCTATCATCCTCTTTTGGCGGCTTTTTCAGCAGATAAATATTA	nem316_all.seq							
10650	AAAGTTCTATCATCCTCTTTTGGCGGCTTTTTCAGCAGATAAATATTA	a909_all.seq							
GGTAGCTTGGTCTAAACAACCTCCAGCAGTTTATTTACACCACCTATTAATC Majority									
10710		10720		10730		10740		10750	
0694	GGTAGCTTGGTCTAAACAACCTCCAGCAGTTTATTTACACCACCTATTAATG	2603_all.seq							
0467	GGTAGCTTGGTCTAAACAACCTCCAGCAGTTTATTTACACCACCTATTAATG	18rs21_all.seq							
0699	GGTAGCTTGGTCTAAACAACCTCCAGCAGTTTATTTACACCACCTATTAATG	cohl_all.seq							
0699	GGTAGCTTGGTCTAAACAACCTCCAGCAGTTTATTTACACCACCTATTAATG	cjb111_all.seq							
0701	GGTAGCTTGGTCTAAACAACCTCCAGCAGTTTATTTACACCACCTATTAATG	nem316_all.seq							
0700	GGTAGCTTGGTCTAAACAACCTCCAGCAGTTTATTTACACCACCTATTAATG	a909_all.seq							
CAAACCTACACCACCTCAGATTCAAGCTATTGGGACCAACGATTAAAGTCACAA Majority									
10760		10770		10780		10790		10800	
0744	CAAACCTACACCACCTCAGATTCAAGCTATTGGGACCAACGATTAAAGTCACAA	2603_all.seq							
0517	CAAACCTACACCACCTCAGATTCAAGCTATTGGGACCAACGATTAAAGTCACAA	18rs21_all.seq							
0749	CAAACCTACACCACCTCAGATTCAAGCTATTGGGACCAACGATTAAAGTCACAA	cohl_all.seq							
0749	CAAACCTACACCACCTCAGATTCAAGCTATTGGGACCAACGATTAAAGTCACAA	cjb111_all.seq							
0751	CAAACCTACACCACCTCAGATTCAAGCTATTGGGACCAACGATTAAAGTCACAA	nem316_all.seq							
0750	CAAACCTACACCACCTCAGATTCAAGCTATTGGGACCAACGATTAAAGTCACAA	a909_all.seq							
ATTCCGGGAATCCGATTTTTCACGGCTTACCGATAAAAAGAGCAGGAAGTTTCAG Majority									
10810		10820		10830		10840		10850	
0794	ATTCCGGGAATCCGATTTTTCACGGCTTACCGATAAAAAGAGCAGGAAGTTTCAG	2603_all.seq							
0567	ATTCCGGGAATCCGATTTTTCACGGCTTACCGATAAAAAGAGCAGGAAGTTTCAG	18rs21_all.seq							
0799	ATTCCGGGAATCCGATTTTTCACGGCTTACCGATAAAAAGAGCAGGAAGTTTCAG	cohl_all.seq							
0799	ATTCCGGGAATCCGATTTTTCACGGCTTACCGATAAAAAGAGCAGGAAGTTTCAG	cjb111_all.seq							
0801	ATTCCGGGAATCCGATTTTTCACGGCTTACCGATAAAAAGAGCAGGAAGTTTCAG	nem316_all.seq							
0800	ATTCCGGGAATCCGATTTTTCACGGCTTACCGATAAAAAGAGCAGGAAGTTTCAG	a909_all.seq							

FIGURE 18 AD

49/487

C A T T A A C A A G A T T G A C C A A G C T A A A G A A G C C T T A G T A G G T G C G A C C T T C A Majority									
10860		10870		10880		10890		10900	
10844	C	A	T	T	A	A	C	A	A
10617	C	A	T	T	A	A	C	A	A
10849	C	A	T	T	A	A	C	A	A
10849	C	A	T	T	A	A	C	A	A
10851	C	A	T	T	A	A	C	A	A
10850	C	A	T	T	A	A	C	A	A
C C T T G T C T A A A C G C A C A A C A G T A G C G G C A G A T C A T C A A G T A C A A G G A G A T Majority									
10910		10920		10930		10940		10950	
10894	C	C	T	T	G	T	C	T	A
10667	C	C	T	T	G	T	C	T	A
10899	C	C	T	T	G	T	C	T	A
10899	C	C	T	T	G	T	C	T	A
10901	C	C	T	T	G	T	C	T	A
10900	C	C	T	T	G	T	C	T	A
T T C A T T C C T G T C A G C A A A G A G A C G A C A G T C G G T C G G A C A A C C C T T A C C T T Majority									
10960		10970		10980		10990		11000	
10944	T	T	C	A	T	T	C	C	T
10717	T	T	C	A	T	T	C	C	T
10949	T	T	C	A	T	T	C	C	T
10949	T	T	C	A	T	T	C	C	T
10951	T	T	C	A	T	T	C	C	T
10950	T	T	C	A	T	T	C	C	T
T G A C A A C C T T A A A C C T G G A T T T T A T G A C C T T A A A G A A A C G A A A G C G C C G A Majority									
11010		11020		11030		11040		11050	
0994	T	G	A	C	A	A	C	C	T
0767	T	G	A	C	A	A	C	C	T
0999	T	G	A	C	A	A	C	C	T
0999	T	G	A	C	A	A	C	C	T
1001	T	G	A	C	A	A	C	C	T
1000	T	G	A	C	A	A	C	C	T
A T G C T T A C G T A C T T G A T C C T A A G A C T T A T G T T G T G G T C G T T C A A A A T T C A Majority									
11060		11070		11080		11090		11100	
1044	A	T	G	C	T	T	A	C	G
3817	A	T	G	C	T	T	A	C	G
1049	A	T	G	C	T	T	A	C	G
1049	A	T	G	C	T	T	A	C	G
1051	A	T	G	C	T	T	A	C	G
1050	A	T	G	C	T	T	A	C	G
G G A A A A A C G A C A A T T G T G G A T G A A G C T A A C T T C A A A G A G G C T G A T T A C C C Majority									
11110		11120		11130		11140		11150	
1094	G	G	A	A	A	A	A	C	G
1867	G	G	A	A	A	A	A	C	G
1099	G	G	A	A	A	A	A	C	G
1099	G	G	A	A	A	A	A	C	G
1101	G	G	A	A	A	A	A	C	G
1100	G	G	A	A	A	A	A	C	G
A A T G C C T G A T A A T A C C A G C C A T G T G G A G T G C G T A G C C T T G C T A C A A C G A A Majority									
11160		11170		11180		11190		11200	
144	A	A	T	G	C	C	T	G	A
917	A	A	T	G	C	C	T	G	A
149	A	A	T	G	C	C	T	G	A
149	A	A	T	G	C	C	T	G	A
151	A	A	T	G	C	C	T	G	A
150	A	A	T	G	C	C	T	G	A

FIGURE 18 AE

50/487

G C A A A G G G T A A A A A T C C T T T A T T T T A A G C A C T T T T T C A A G C A T T T T G T C T Majority									
11210		11220		11230		11240		11250	
11194	G C A A A G G G T A A A A A T C C T T T A T T T T A A G C A C T T T T T C A A G C A T T T T G T C T	2603_all.seq							
10967	G C A A A G G G T A A A A A T C C T T T A T T T T A A G C A C T T T T T C A A G C A T T T T G T C T	18rs21_all.seq							
11199	G C A A A G G G T A A A A A T C C T T T A T T T T A A G C A C T T T T T C A A G C A T T T T G T C T	coh1_all.seq							
11199	G C A A A G G G T A A A A A T C C T T T A T T T T A A G C A C T T T T T C A A G C A T T T T G T C T	cjb111_all.seq							
11201	G C A A A G G G T A A A A A T C C T T T A T T T T A A G C A C T T T T T C A A G C A T T T T G T C T	nem316_all.seq							
11200	G C A A A G G G T A A A A A T C C T T T A T T T T A A G C A C T T T T T C A A G C A T T T T G T C T	a909_all.seq							
T T A T T G A A A A G A G T G A T T T T A A C A T A A A A A A G C T A T T A A A A A C A T A T T G Majority									
11260		11270		11280		11290		11300	
11244	T T A T T G A A A A G A G T G A T T T T A A C A T A A A A A A G C T A T T A A A A A A C A T A T T G	2603_all.seq							
11017	T T A T T G A A A A G A G T G A T T T T A A C A T A A A A A A G C T A T T A A A A A A C A T A T T G	18rs21_all.seq							
11249	T T A T T G A A A A G A G T G A T T T T A A C A T A A A A A A G C T A T T A A A A A A C A T A T T G	coh1_all.seq							
11249	T T A T T G A A A A G A G T G A T T T T A A C A T A A A A A A G C T A T T A A A A A A C A T A T T G	cjb111_all.seq							
11251	T T A T T G A A A A G A G T G A T T T T A A C A T A A A A A A G C T A T T A A A A A A C A T A T T G	nem316_all.seq							
11250	T T A T T G A A A A G A G T G A T T T T A A C A T A A A A A A G C T A T T A A A A A A C A T A T T G	a909_all.seq							
A C C T G A C C G T T T G T T T T G A A G T G G C T T G C G T A G A C A A A A A A A T A G A T A C G Majority									
11310		11320		11330		11340		11350	
11294	A C C T G A C C G T T T G T T T T G A A G T G G C T T G C G T A G A C A A A A A A A T A G A T A C G	2603_all.seq							
11067	A C C T G A C C G T T T G T T T T G A A G T G G C T T G C G T A G A C A A A A A A A T A G A T A C G	18rs21_all.seq							
11299	A C C T G A C C G T T T G T T T T G A A G T G G C T T G C G T A G A C A A A A A A A T A G A T A C G	coh1_all.seq							
11299	A C C T G A C C G T T T G T T T T G A A G T G G C T T G C G T A G A C A A A A A A A T A G A T A C G	cjb111_all.seq							
11301	A C C T G A C C G T T T G T T T T G A A G T G G C T T G C G T A G A C A A A A A A A T A G A T A C G	nem316_all.seq							
11300	A C C T G A C C G T T T G T T T T G A A G T G G C T T G C G T A G A C A A A A A A A T A G A T A C G	a909_all.seq							
T C A G A T A A A T T T C T G G C A T T A C G A G A A C A T T T T T A G A G T G T T C T C T T T T T Majority									
11360		11370		11380		11390		11400	
11344	T C A G A T A A A T T T C T G G C A T T A C G A G A A C A T T T T T A G A G T G T T C T C T T T T T	2603_all.seq							
11117	T C A G A T A A A T T T C T G G C A T T A C G A G A A C A T T T T T A G A G T G T T C T C T T T T T	18rs21_all.seq							
11349	T C A G A T A A A T T T C T G G C A T T A C G A G A A C A T T T T T A G A G T G T T C T C T T T T T	coh1_all.seq							
11349	T C A G A T A A A T T T C T G G C A T T A C G A G A A C A T T T T T A G A G T G T T C T C T T T T T	cjb111_all.seq							
11351	T C A G A T A A A T T T C T G G C A T T A C G A G A A C A T T T T T A G A G T G T T C T C T T T T T	nem316_all.seq							
11350	T C A G A T A A A T T T C T G G C A T T A C G A G A A C A T T T T T A G A G T G T T C T C T T T T T	a909_all.seq							
T T A G T T T A C G G A G G A A A A A T A T A T A T G G A A A A A C A G G A T T C A C G A G T T C T Majority									
11410		11420		11430		11440		11450	
11394	T T A G T T T A C G G A G G A A A A A T A T A T A T G G A A A A A C A G G A T T C A C G A G T T C T	2603_all.seq							
11167	T T A G T T T A C G G A G G A A A A A T A T A T A T G G A A A A A C A G G A T T C A C G A G T T C T	18rs21_all.seq							
11399	T T A G T T T A C G G A G G A A A A A T A T A T A T G G A A A A A C A G G A T T C A C G A G T T C T	coh1_all.seq							
11399	T T A G T T T A C G G A G G A A A A A T A T A T A T G G A A A A A C A G G A T T C A C G A G T T C T	cjb111_all.seq							
1401	T T A G T T T A C G G A G G A A A A A T A T A T A T G G A A A A A C A G G A T T C A C G A G T T C T	nem316_all.seq							
1400	T T A G T T T A C G G A G G A A A A A T A T A T A T G G A A A A A C A G G A T T C A C G A G T T C T	a909_all.seq							
C A T C C A T T G G G A G G G G A A T T C T G G G G A C A A G C T C A T T G A A C A C C A A A C C A Majority									
11460		11470		11480		11490		11500	
1444	C A T C C A T T G G G A G G G G A A T T C T G G G G A C A A G C T C A T T G A A C A C C A A A C C A	2603_all.seq							
1217	C A T C C A T T G G G A G G G G A A T T C T G G G G A C A A G C T C A T T G A A C A C C A A A C C A	18rs21_all.seq							
1449	C A T C C A T T G G G A G G G G A A T T C T G G G G A C A A G C T C A T T G A A C A C C A A A C C A	coh1_all.seq							
1449	C A T C C A T T G G G A G G G G A A T T C T G G G G A C A A G C T C A T T G A A C A C C A A A C C A	cjb111_all.seq							
1451	C A T C C A T T G G G A G G G G A A T T C T G G G G A C A A G C T C A T T G A A C A C C A A A C C A	nem316_all.seq							
1450	C A T C C A T T G G G A G G G G A A T T C T G G G G A C A A G C T C A T T G A A C A C C A A A C C A	a909_all.seq							
G C C C A A C G G G G T G G T A C T A C C A A G T C G A T C G T A G C T T T A G T C A A C C A A A A Majority									
11510		11520		11530		11540		11550	
1494	G C C C A A C G G G G T G G T A C T A C C A A G T C G A T C G T A G C T T T A G T C A A C C A A A A	2603_all.seq							
1267	G C C C A A C G G G G T G G T A C T A C C A A G T C G A T C G T A G C T T T A G T C A A C C A A A A	18rs21_all.seq							
1499	G C C C A A C G G G G T G G T A C T A C C A A G T C G A T C G T A G C T T T A G T C A A C C A A A A	coh1_all.seq							
1499	G C C C A A C G G G G T G G T A C T A C C A A G T C G A T C G T A G C T T T A G T C A A C C A A A A	cjb111_all.seq							
1501	G C C C A A C G G G G T G G T A C T A C C A A G T C G A T C G T A G C T T T A G T C A A C C A A A A	nem316_all.seq							
1500	G C C C A A C G G G G T G G T A C T A C C A A G T C G A T C G T A G C T T T A G T C A A C C A A A A	a909_all.seq							

FIGURE 18 AF

51/487

G G C G A A C C G C C C A G A A T G A T C G A A A G G C A C T A G A A A G T G T C C G T A A T G A T Majority										
11560		11570		11580		11590		11600		
11544	G G C G A A C C G C C C A G A A T G A T C C A A A G G C A C T A G A A A G T G T C C G T A A T G A T									2603_all.seq
11317	G G C G A A C C G C C C A G A A T G A T C C A A A G G C A C T A G A A A G T G T C C G T A A T G A T									18rs21_all.seq
11549	G G C G A A C C G C C C A G A A T G A T C C A A A G G C A C T A G A A A G T G T C C G T A A T G A T									cohl1_all.seq
11551	G G C G A A C C G C C C A G A A T G A T C C A A A G G C A C T A G A A A G T G T C C G T A A T G A T									cjb111_all.seq
11550	G G C G A A C C G C C C A G A A T G A T C C A A A G G C A C T A G A A A G T G T C C G T A A T G A T									nem316_all.seq
										a909_all.seq
T C G A T T T C G G G C G G T G A T G A T G T C A T G G G T T A T G C T T A T A G C A A A T G T A C Majority										
11610		11620		11630		11640		11650		
11594	T C G A T T T C G G G C G G T G A T G A T G T C A T G G G T T A T G C T T A T A G C A A A T G T A C									2603_all.seq
11367	T C G A T T T C G G G C G G T G A T G A T G T C A T G G G T T A T G C T T A T A G C A A A T G T A C									18rs21_all.seq
11599	T C G A T T T C G G G C G G T G A T G A T G T C A T G G G T T A T G C T T A T A G C A A A T G T A C									cohl1_all.seq
11599	T C G A T T T C G G G C G G T G A T G A T G T C A T G G G T T A T G C T T A T A G C A A A T G T A C									cjb111_all.seq
11601	T C G A T T T C G G G C G G T G A T G A T G T C A T G G G T T A T G C T T A T A G C A A A T G T A C									nem316_all.seq
11600	T C G A T T T C G G G C G G T G A T G A T G T C A T G G G T T A T G C T T A T A G C A A A T G T A C									a909_all.seq
T T G G G G A G T T T G C G G C A C G A A T T A A T C A G T G G G A C T G A A A C T C A A A G C G T T G Majority										
11660		11670		11680		11690		11700		
1644	T T G G G G A G T T T G C G G C A C G A A T T A A T C A G T G G G A C T G A A A C T C A A A G C G T T G									2603_all.seq
1417	T T G G G G A G T T T G C G G C A C G A A T T A A T C A G T G G G A C T G A A A C T C A A A G C G T T G									18rs21_all.seq
1649	T T G G G G A G T T T G C G G C A C G A A T T A A T C A G T G G G A C T G A A A C T C A A A G C G T T G									cohl1_all.seq
1649	T T G G G G A G T T T G C G G C A C G A A T T A A T C A G T G G G A C T G A A A C T C A A A G C G T T G									cjb111_all.seq
1651	T T G G G G A G T T T G C G G C A C G A A T T A A T C A G T G G G A C T G A A A C T C A A A G C G T T G									nem316_all.seq
1650	T T G G G G A G T T T G C G G C A C G A A T T A A T C A G T G G G A C T G A A A C T C A A A G C G T T G									a909_all.seq
A A A T G G T G A G A A G A T T A C C A T T A C C A G T T C A A T G G G A A A T G G T C A G G A T T Majority										
11710		11720		11730		11740		11750		
1694	A A A T G G T G A G A A G A T T A C C A T T A C C A G T T C A A T G G G A A A T G G T C A G G A T T									2603_all.seq
1467	A A A T G G T G A G A A G A T T A C C A T T A C C A G T T C A A T G G G A A A T G G T C A G G A T T									18rs21_all.seq
1699	A A A T G G T G A G A A G A T T A C C A T T A C C A G T T C A A T G G G A A A T G G T C A G G A T T									cohl1_all.seq
1699	A A A T G G T G A G A A G A T T A C C A T T A C C A G T T C A A T G G G A A A T G G T C A G G A T T									cjb111_all.seq
1701	A A A T G G T G A G A A G A T T A C C A T T A C C A G T T C A A T G G G A A A T G G T C A G G A T T									nem316_all.seq
1700	A A A T G G T G A G A A G A T T A C C A T T A C C A G T T C A A T G G G A A A T G G T C A G G A T T									a909_all.seq
G G G T T G G A A C A G C C G A A A G A C T A G A T G G T G A A A C T G A T A C A G T T C C A A A A Majority										
11760		11770		11780		11790		11800		
1744	G G G T T G G A A C A G C C G A A A G A C T A G A T G G T G A A A C T G A T A C A G T T C C A A A A									2603_all.seq
1517	G G G T T G G A A C A G C C G A A A G A C T A G A T G G T G A A A C T G A T A C A G T T C C A A A A									18rs21_all.seq
1749	G G G T T G G A A C A G C C G A A A G A C T A G A T G G T G A A A C T G A T A C A G T T C C A A A A									cohl1_all.seq
1749	G G G T T G G A A C A G C C G A A A G A C T A G A T G G T G A A A C T G A T A C A G T T C C A A A A									cjb111_all.seq
1751	G G G T T G G A A C A G C C G A A A G A C T A G A T G G T G A A A C T G A T A C A G T T C C A A A A									nem316_all.seq
1750	G G G T T G G A A C A G C C G A A A G A C T A G A T G G T G A A A C T G A T A C A G T T C C A A A A									a909_all.seq
G A A G C T A C T A T T C T C T C T T T T A G G A A A C T A G T T A T G C T T C G T A T A T A G C Majority										
11810		11820		11830		11840		11850		
794	G A A G C T A C T A T T C T C T C T T T T A G G A A A C T A G T T A T G C T T C G T A T A T A G C									2603_all.seq
567	G A A G C T A C T A T T C T C T C T T T T A G G A A A C T A G T T A T G C T T C G T A T A T A G C									18rs21_all.seq
799	G A A G C T A C T A T T C T C T C T T T T A G G A A A C T A G T T A T G C T T C G T A T A T A G C									cohl1_all.seq
799	G A A G C T A C T A T T C T C T C T T T T A G G A A A C T A G T T A T G C T T C G T A T A T A G C									cjb111_all.seq
801	G A A G C T A C T A T T C T C T C T T T T A G G A A A C T A G T T A T G C T T C G T A T A T A G C									nem316_all.seq
800	G A A G C T A C T A T T C T C T C T T T T A G G A A A C T A G T T A T G C T T C G T A T A T A G C									a909_all.seq
C T A C G G A A C T A T A T C T T T C G T C A C A T T A C A T C T A C A G A T A G T A C C A T G A A Majority										
11860		11870		11880		11890		11900		
844	C T A C G G A A C T A T A T C T T T C G T C A C A T T A C A T C T A C A G A T A G T A C C A T G A A									2603_all.seq
617	C T A C G G A A C T A T A T C T T T C G T C A C A T T A C A T C T A C A G A T A G T A C C A T G A A									18rs21_all.seq
849	C T A C G G A A C T A T A T C T T T C G T C A C A T T A C A T C T A C A G A T A G T A C C A T G A A									cohl1_all.seq
849	C T A C G G A A C T A T A T C T T T C G T C A C A T T A C A T C T A C A G A T A G T A C C A T G A A									cjb111_all.seq
851	C T A C G G A A C T A T A T C T T T C G T C A C A T T A C A T C T A C A G A T A G T A C C A T G A A									nem316_all.seq
850	C T A C G G A A C T A T A T C T T T C G T C A C A T T A C A T C T A C A G A T A G T A C C A T G A A									a909_all.seq

FIGURE 18 AG

TTTTGCTTATATGACCAAGTAAAGTGAGGATATACTAACAAATGAAATAT										Majority
11910		11920		11930		11940		11950		
11894	TTTTGCTTATATGACCAAGTAAAGTGAGGATATACTAACAAATGAAATAT									2603_all.seq
11667	TTTTGCTTATATGACCAAGTAAAGTGAGGATATACTAACAAATGAAATAT									18rs21_all.seq
11899	TTTTGCTTATATGACCAAGTAAAGTGAGGATATACTAACAAATGAAATAT									cohl_all.seq
11899	TTTTGCTTATATGACCAAGTAAAGTGAGGATATACTAACAAATGAAATAT									cjb111_all.seq
11901	TTTTGCTTATATGACCAAGTAAAGTGAGGATATACTAACAAATGAAATAT									nen316_all.seq
11900	TTTTGCTTATATGACCAAGTAAAGTGAGGATATACTAACAAATGAAATAT									a909_all.seq
TTATTATCGTATTTGTCCATTTTATCGAAAAGTTTG-CATATTATCATTAT										Majority
11960		11970		11980		11990		12000		
11944	TTATTATCGTATTTGTCCATTTTATCGAAAAGTTTG-CATATTATCATTAT									2603_all.seq
11717	TTATTATCGTATTTGTCCATTTTATCGAAAAGTTTG-CATATTATCATTAT									18rs21_all.seq
11949	TTATTATCGTATTTGTCCATTTTATCGAAAAGTTTG-CATATTATCATTAT									cohl_all.seq
11949	TTATTATCGTATTTGTCCATTTTATCGAAAAGTTTG-CATATTATCATTAT									cjb111_all.seq
11951	TTATTATCGTATTTGTCCATTTTATCGAAAAGTTTG-CATATTATCATTAT									nen316_all.seq
11950	TTATTATCGTATTTGTCCATTTTATCGAAAAGTTTG-CATATTATCATTAT									a909_all.seq
GTTTGATAAAGATGCAAAATATAATGATAGTAGGAGCTAAATATGGATATTT										Majority
12010		12020		12030		12040		12050		
11994	GTTTGATAAAGATGCAAAATATAATGATAGTAGGAGCTAAATATGGATATTT									2603_all.seq
11767	GTTTGATAAAGATGCAAAATATAATGATAGTAGGAGCTAAATATGGATATTT									18rs21_all.seq
11999	GTTTGATAAAGATGCAAAATATAATGATAGTAGGAGCTAAATATGGATATTT									cohl_all.seq
11999	GTTTGATAAAGATGCAAAATATAATGATAGTAGGAGCTAAATATGGATATTT									cjb111_all.seq
12001	GTTTGATAAAGATGCAAAATATAATGATAGTAGGAGCTAAATATGGATATTT									nen316_all.seq
12000	GTTTGATAAAGATGCAAAATATAATGATAGTAGGAGCTAAATATGGATATTT									a909_all.seq
AAAAAATCAAGAGTATCCTAAGTGCTTTTCCATTTTGA AATTCA AATATAG										Majority
12060		12070		12080		12090		12100		
12044	AAAAAATCAAGAGTATCCTAAGTGCTTTTCCATTTTGA AATTCA AATATAG									2603_all.seq
11817	AAAAAATCAAGAGTATCCTAAGTGCTTTTCCATTTTGA AATTCA AATATAG									18rs21_all.seq
12049	AAAAAATCAAGAGTATCCTAAGTGCTTTTCCATTTTGA AATTCA AATATAG									cohl_all.seq
12049	AAAAAATCAAGAGTATCCTAAGTGCTTTTCCATTTTGA AATTCA AATATAG									cjb111_all.seq
12051	AAAAAATCAAGAGTATCCTAAGTGCTTTTCCATTTTGA AATTCA AATATAG									nen316_all.seq
12050	AAAAAATCAAGAGTATCCTAAGTGCTTTTCCATTTTGA AATTCA AATATAG									a909_all.seq
CTAATAGTTCTAGAACTTCTAATTGTTTTTCGTCGACGATATGAATTTTC										Majority
12110		12120		12130		12140		12150		
2094	CTAATAGTTCTAGAACTTCTAATTGTTTTTCGTCGACGATATGAATTTTC									2603_all.seq
1867	CTAATAGTTCTAGAACTTCTAATTGTTTTTCGTCGACGATATGAATTTTC									18rs21_all.seq
2099	CTAATAGTTCTAGAACTTCTAATTGTTTTTCGTCGACGATATGAATTTTC									cohl_all.seq
2099	CTAATAGTTCTAGAACTTCTAATTGTTTTTCGTCGACGATATGAATTTTC									cjb111_all.seq
2101	CTAATAGTTCTAGAACTTCTAATTGTTTTTCGTCGACGATATGAATTTTC									nen316_all.seq
2100	CTAATAGTTCTAGAACTTCTAATTGTTTTTCGTCGACGATATGAATTTTC									a909_all.seq
AATCTTAACTGTTAGGATTCCACCTCCCTTTGCTTAAAGAAAAAAGGTC A										Majority
12160		12170		12180		12190		12200		
2144	AATCTTAACTGTTAGGATTCCACCTCCCTTTGCTTAAAGAAAAAAGGTC A									2603_all.seq
1917	AATCTTAACTGTTAGGATTCCACCTCCCTTTGCTTAAAGAAAAAAGGTC A									18rs21_all.seq
2149	AATCTTAACTGTTAGGATTCCACCTCCCTTTGCTTAAAGAAAAAAGGTC A									cohl_all.seq
2149	AATCTTAACTGTTAGGATTCCACCTCCCTTTGCTTAAAGAAAAAAGGTC A									cjb111_all.seq
2151	AATCTTAACTGTTAGGATTCCACCTCCCTTTGCTTAAAGAAAAAAGGTC A									nen316_all.seq
2150	AATCTTAACTGTTAGGATTCCACCTCCCTTTGCTTAAAGAAAAAAGGTC A									a909_all.seq
GGTCGTTTTAGATAA ACTTTGTCAAACAAGCTCAAGCTATCTAAAAATAGTT										Majority
12210		12220		12230		12240		12250		
2194	GGTCGTTTTAGATAA ACTTTGTCAAACAAGCTCAAGCTATCTAAAAATAGTT									2603_all.seq
1967	GGTCGTTTTAGATAA ACTTTGTCAAACAAGCTCAAGCTATCTAAAAATAGTT									18rs21_all.seq
2199	GGTCGTTTTAGATAA ACTTTGTCAAACAAGCTCAAGCTATCTAAAAATAGTT									cohl_all.seq
2199	GGTCGTTTTAGATAA ACTTTGTCAAACAAGCTCAAGCTATCTAAAAATAGTT									cjb111_all.seq
2201	GGTCGTTTTAGATAA ACTTTGTCAAACAAGCTCAAGCTATCTAAAAATAGTT									nen316_all.seq
2200	GGTCGTTTTAGATAA ACTTTGTCAAACAAGCTCAAGCTATCTAAAAATAGTT									a909_all.seq

FIGURE 18 AH

53/487

T G A A A T G G G C A T T A C T C T A G T T T T T A A T A A G C T A T C T G A T G A G C A G A A G C Majority									
12260		12270		12280		12290		12300	
12244	T	G	A	A	T	G	G	G	C
12017	T	G	A	A	T	G	G	G	C
12249	T	G	A	A	T	G	G	G	C
12249	T	G	A	A	T	G	G	G	C
12251	T	G	A	A	T	G	G	G	C
12250	T	G	A	A	T	G	G	G	C
A G A A G T T A A T G C A T G T T G G G A A G T C T T A T T T T G A C T A T C A A G A A A A T G C T Majority									
12310		12320		12330		12340		12350	
12294	A	G	A	A	G	T	T	A	A
12067	A	G	A	A	G	T	T	A	A
12299	A	G	A	A	G	T	T	A	A
12299	A	G	A	A	G	T	T	A	A
12301	A	G	A	A	G	T	T	A	A
12300	A	G	A	A	G	T	T	A	A
C T T A T C C C A C A A T T A G G T T T T C T A T A T T C T A A A T T A A C T A A A A A A A T T G A Majority									
12360		12370		12380		12390		12400	
2344	C	T	T	A	T	C	C	C	A
2117	C	T	T	A	T	C	C	C	A
2349	C	T	T	A	T	C	C	C	A
2349	C	T	T	A	T	C	C	C	A
2351	C	T	T	A	T	C	C	C	A
2350	C	T	T	A	T	C	C	C	A
A C T T G A T A A T C G G T T G T C T C C G A C T G A A C A A A A G T T A T T G A T T A C C T T A T Majority									
12410		12420		12430		12440		12450	
2394	A	C	T	T	G	A	T	A	A
2167	A	C	T	T	G	A	T	A	A
2399	A	C	T	T	G	A	T	A	A
2399	A	C	T	T	G	A	T	A	A
2401	A	C	T	T	G	A	T	A	A
2400	A	C	T	T	G	A	T	A	A
T A T T A C A T A C T A A A G G T T T A A T C A T T G A T A T G T A A G A A G T A A G T C A G C T A Majority									
12460		12470		12480		12490		12500	
2444	T	A	T	T	A	C	A	T	A
2217	T	A	T	T	A	C	A	T	A
2449	T	A	T	T	A	C	A	T	A
2449	T	A	T	T	A	C	A	T	A
2451	T	A	T	T	A	C	A	T	A
2450	T	A	T	T	A	C	A	T	A
A C C G A T C T T T C T A T T C T A A A A C T T A T A T T G T T G C T T T A G A A A T T T T A A A G Majority									
12510		12520		12530		12540		12550	
2494	A	C	C	G	A	T	C	T	T
2267	A	C	C	G	A	T	C	T	T
2499	A	C	C	G	A	T	C	T	T
2499	A	C	C	G	A	T	C	T	T
2501	A	C	C	G	A	T	C	T	T
2500	A	C	C	G	A	T	C	T	T
A G C G T G C A T G G C T T C A T A A T A A A C A G A A A T C T T A C C A A T T T G C G A A G C C A Majority									
12560		12570		12580		12590		12600	
2544	A	G	C	G	T	G	C	A	T
2317	A	G	C	G	T	G	C	A	T
2549	A	G	C	G	T	G	C	A	T
2549	A	G	C	G	T	G	C	A	T
2551	A	G	C	G	T	G	C	A	T
2550	A	G	C	G	T	G	C	A	T

FIGURE 18 AI

54/487

WO 2006/078318

PCT/US2005/027239

Page 37

A A A A A T A T G C A T A T T T G A A G A G T C C A A A G A T C T A A T A G A T A G T C C A G T T A G Majority									
12610		12620		12630		12640		12650	
2594	A A A A A T A T G A T A T T T G A A G A G T C C A A A G A T C T A A T A G A T A G T C C A G T T A G								2603_all.seq
2367	A A A A A T A T G A T A T T T G A A G A G T C C A A A G A T C T A A T A G A T A G T C C A G T T A G								18rs21_all.seq
2599	A A A A A T A T G A T A T T T G A A G A G T C C A A A G A T C T A A T A G A T A G T C C A G T T A G								cohl_all.seq
2599	A A A A A T A T G A T A T T T G A A G A G T C C A A A G A T C T A A T A G A T A G T C C A G T T A G								cjb111_all.seq
2601	A A A A A T A T G A T A T T T G A A G A G T C C A A A G A T C T A A T A G A T A G T C C A G T T A G								nem316_all.seq
2600	A A A A A T A T G A T A T T T G A A G A G T C C A A A G A T C T A A T A G A T A G T C C A G T T A G								a909_all.seq
A G A A G C C G T T G A T T A T A A C T G A T A A G G A T T T T C A A A A A T T A A A A C A A G A G C Majority									
12660		12670		12680		12690		12700	
2644	A G A A G C C G T T G A T T A T A A G T G A T A A G G A T T T T C A A A A A T T A A A A C A A G A G C								2603_all.seq
2417	A G A A G C C G T T G A T T A T A A G T G A T A A G G A T T T T C A A A A A T T A A A A C A A G A G C								18rs21_all.seq
2649	A G A A G C C G T T G A T T A T A A G T G A T A A G G A T T T T C A A A A A T T A A A A C A A G A G C								cohl_all.seq
2649	A G A A G C C G T T G A T T A T A A G T G A T A A G G A T T T T C A A A A A T T A A A A C A A G A G C								cjb111_all.seq
2651	A G A A G C C G T T G A T T A T A A G T G A T A A G G A T T T T C A A A A A T T A A A A C A A G A G C								nem316_all.seq
2650	A G A A G C C G T T G A T T A T A A G T G A T A A G G A T T T T C A A A A A T T A A A A C A A G A G C								a909_all.seq
T A T T A T T T T A A C C G A C T T A T T T T A A A G A C T T A T C A T A T C T A G G C T T G C T T Majority									
12710		12720		12730		12740		12750	
2694	T A T T A T T T T A A C C G A C T T A T T T T A A A G A C T T A T C A T A T C T A G G C T T G C T T								2603_all.seq
2467	T A T T A T T T T A A C C G A C T T A T T T T A A A G A C T T A T C A T A T C T A G G C T T G C T T								18rs21_all.seq
2699	T A T T A T T T T A A C C G A C T T A T T T T A A A G A C T T A T C A T A T C T A G G C T T G C T T								cohl_all.seq
2699	T A T T A T T T T A A C C G A C T T A T T T T A A A G A C T T A T C A T A T C T A G G C T T G C T T								cjb111_all.seq
2701	T A T T A T T T T A A C C G A C T T A T T T T A A A G A C T T A T C A T A T C T A G G C T T G C T T								nem316_all.seq
2700	T A T T A T T T T A A C C G A C T T A T T T T A A A G A C T T A T C A T A T C T A G G C T T G C T T								a909_all.seq
G A T G A T T C G G A A A A A T A C G G A G A C T A T A C T A T T T T C A A G C A A A A G A T A C A A Majority									
12760		12770		12780		12790		12800	
2744	G A T G A T T C G G A A A A A T A C G G A G A C T A T A C T A T T T T C A A G C A A A A G A T A C A A								2603_all.seq
2517	G A T G A T T C G G A A A A A T A C G G A G A C T A T A C T A T T T T C A A G C A A A A G A T A C A A								18rs21_all.seq
2749	G A T G A T T C G G A A A A A T A C G G A G A C T A T A C T A T T T T C A A G C A A A A G A T A C A A								cohl_all.seq
2749	G A T G A T T C G G A A A A A T A C G G A G A C T A T A C T A T T T T C A A G C A A A A G A T A C A A								cjb111_all.seq
2751	G A T G A T T C G G A A A A A T A C G G A G A C T A T A C T A T T T T C A A G C A A A A G A T A C A A								nem316_all.seq
2750	G A T G A T T C G G A A A A A T A C G G A G A C T A T A C T A T T T T C A A G C A A A A G A T A C A A								a909_all.seq
A A G T T T C G A A T C A A G T C T T C A A C T A T A C A T C C T T C A A A G T C A T C G G C T A G Majority									
12810		12820		12830		12840		12850	
2794	A A G T T T C G A A T C A A G T C T T C A A C T A T A C A T C C T T C A A A G T C A T C G G C T A G								2603_all.seq
2567	A A G T T T C G A A T C A A G T C T T C A A C T A T A C A T C C T T C A A A G T C A T C G G C T A G								18rs21_all.seq
2799	A A G T T T C G A A T C A A G T C T T C A A C T A T A C A T C C T T C A A A G T C A T C G G C T A G								cohl_all.seq
2799	A A G T T T C G A A T C A A G T C T T C A A C T A T A C A T C C T T C A A A G T C A T C G G C T A G								cjb111_all.seq
2801	A A G T T T C G A A T C A A G T C T T C A A C T A T A C A T C C T T C A A A G T C A T C G G C T A G								nem316_all.seq
2800	A A G T T T C G A A T C A A G T C T T C A A C T A T A C A T C C T T C A A A G T C A T C G G C T A G								a909_all.seq
A G A T T T G G A A T T A T G A A C C A A T C C C T T T G A T T A C T A G A A A A A T A A A T A G C Majority									
12860		12870		12880		12890		12900	
2844	A G A T T T G G A A T T A T G A A C C A A T C C C T T T G A T T A C T A G A A A A A T A A A T A G C								2603_all.seq
2617	A G A T T T G G A A T T A T G A A C C A A T C C C T T T G A T T A C T A G A A A A A T A A A T A G C								18rs21_all.seq
2849	A G A T T T G G A A T T A T G A A C C A A T C C C T T T G A T T A C T A G A A A A A T A A A T A G C								cohl_all.seq
2849	A G A T T T G G A A T T A T G A A C C A A T C C C T T T G A T T A C T A G A A A A A T A A A T A G C								cjb111_all.seq
2851	A G A T T T G G A A T T A T G A A C C A A T C C C T T T G A T T A C T A G A A A A A T A A A T A G C								nem316_all.seq
2850	A G A T T T G G A A T T A T G A A C C A A T C C C T T T G A T T A C T A G A A A A A T A A A T A G C								a909_all.seq
T T G C A G A G G C T A A C G T G A C A C T G G T T G A T C C A A T C T C G C T T T A T T T A A C A Majority									
12910		12920		12930		12940		12950	
2894	T T G C A G A G G C T A A C G T G A C A C T G G T T G A T C C A A T C T C G C T T T A T T T A A C A								2603_all.seq
2667	T T G C A G A G G C T A A C G T G A C A C T G G T T G A T C C A A T C T C G C T T T A T T T A A C A								18rs21_all.seq
2899	T T G C A G A G G C T A A C G T G A C A C T G G T T G A T C C A A T C T C G C T T T A T T T A A C A								cohl_all.seq
2899	T T G C A G A G G C T A A C G T G A C A C T G G T T G A T C C A A T C T C G C T T T A T T T A A C A								cjb111_all.seq
2901	T T G C A G A G G C T A A C G T G A C A C T G G T T G A T C C A A T C T C G C T T T A T T T A A C A								nem316_all.seq
2900	T T G C A G A G G C T A A C G T G A C A C T G G T T G A T C C A A T C T C G C T T T A T T T A A C A								a909_all.seq

FIGURE 18 AJ

55/487

WO 2006/078318

Alignment Report of A1-1 alignment, using J. Hein method with Weighted residue weight table.
Thursday, July 29, 2004 5:46 PM.

PCT/US2005/027239

Page 38

CTAAGAATGATGAAGACCCCTCGTATTGAAGAAGAAGTTGAGCAGCTAGAA Majority
12960 12970 12980 12990 13000

12944 CTAAGAATGATGAAGACCCCTCGTATTGAAGAAGAAGTTGAGCAGCTAGAA 2603_all.seq
12717 CTAAGAATGATGAAGACCCCTCGTATTGAAGAAGAAGTTGAGCAGCTAGAA 18rs21_all.seq
12949 CTAAGAATGATGAAGACCCCTCGTATTGAAGAAGAAGTTGAGCAGCTAGAA coh1_all.seq
12949 CTAAGAATGATGAAGACCCCTCGTATTGAAGAAGAAGTTGAGCAGCTAGAA cjb111_all.seq
12951 CTAAGAATGATGAAGACCCCTCGTATTGAAGAAGAAGTTGAGCAGCTAGAA nem316_all.seq
12950 CTAAGAATGATGAAGACCCCTCGTATTGAAGAAGAAGTTGAGCAGCTAGAA a909_all.seq

CATAAGAT

Majority

12994 CATAAGAT
12767 CATAAGAT
12999 CATAAGAT
12999 CA
13000 A
13000 CATAAGAT

2603_all.seq
18rs21_all.seq
coh1_all.seq
cjb111_all.seq
nem316_all.seq
a909_all.seq

Decoration *Decoration #1*: Shade (with solid black) residues that differ from the Consensus.

FIGURE 18 AK

Figure 19

57/487

C T T T T A T T T A C T T C A C T T T C T T A A C C A T C C T T G G C T A A A A A G A T A T A C Majority									
	460	470	480	490	500				
451	C T T T T A T T T A C T T C A C T T T C T T A A C C A A T C C T T G G C T A A A A A G A T A T A C					2603_a12.seq			
451	C T T T T A T T T A C T T C A C T T T C T T A A C C A A T C C T T G G C T A A A A A G A T A T A C					18rs21_a12.seq			
451	C T T T T A T T T A C T T C A C T T T C T T A A C C A A T C C T T G G C T A A A A A G A T A T A C					515_a12.seq			
451	C T T T T A T T T A C T T C A C T T T C T T A A C C A A T C C T T G G C T A A A A A G A T A T A C					cjb111_a12.seq			
451	C T T T T A T T T A C T T C A C T T T C T T A A C C A A T C C T T G G C T A A A A A G A T A T A C					h36b_a12.seq			
G C A G T T A G A T T C A A A A T A C C A T A A G C A A G T A T A A A A C C A G C T A A A A C A T C Majority									
	510	520	530	540	550				
501	G C A G T T A G A T T C A A A A T A C C A T A A G C A A G T A T A A A A C C A G C T A A A A C A T C					2603_a12.seq			
501	G C A G T T A G A T T C A A A A T A C C A T A A G C A A G T A T A A A A C C A G C T A A A A C A T C					18rs21_a12.seq			
501	G C A G T T A G A T T C A A A A T A C C A T A A G C A A G T A T A A A A C C A G C T A A A A C A T C					515_a12.seq			
501	G C A G T T A G A T T C A A A A T A C C A T A A G C A A G T A T A A A A C C A G C T A A A A C A T C					cjb111_a12.seq			
501	G C A G T T A G A T T C A A A A T A C C A T A A G C A A G T A T A A A A C C A G C T A A A A C A T C					h36b_a12.seq			
T G T C G G A A A A T G A A C C C C T A G G T A A A T A C G A G A T A A C C C A A T T A A A A A A A Majority									
	560	570	580	590	600				
551	T G T C G G A A A A T G A A C C C C T A G G T A A A T A C G A G A T A A C C C A A T T A A A A A A A					2603_a12.seq			
551	T G T C G G A A A A T G A A C C C C T A G G T A A A T A C G A G A T A A C C C A A T T A A A A A A A					18rs21_a12.seq			
551	T G T C G G A A A A T G A A C C C C T A G G T A A A T A C G A G A T A A C C C A A T T A A A A A A A					515_a12.seq			
551	T G T C G G A A A A T G A A C C C C T A G G T A A A T A C G A G A T A A C C C A A T T A A A A A A A					cjb111_a12.seq			
551	T G T C G G A A A A T G A A C C C C T A G G T A A A T A C G A G A T A A C C C A A T T A A A A A A A					h36b_a12.seq			
T G A G C A A A C C C A A A G T A C C T T G G C A C A A C A G T T T C C A T A T A C T C T T A G G C Majority									
	610	620	630	640	650				
601	T G A G C A A A C C C A A A G T A C C T T G G C A C A A C A G T T T C C A T A T A C T C T T A G G C					2603_a12.seq			
601	T G A G C A A A C C C A A A G T A C C T T G G C A C A A C A G T T T C C A T A T A C T C T T A G G C					18rs21_a12.seq			
601	T G A G C A A A C C C A A A G T A C C T T G G C A C A A C A G T T T C C A T A T A C T C T T A G G C					515_a12.seq			
601	T G A G C A A A C C C A A A G T A C C T T G G C A C A A C A G T T T C C A T A T A C T C T T A G G C					cjb111_a12.seq			
601	T G A G C A A A C C C A A A G T A C C T T G G C A C A A C A G T T T C C A T A T A C T C T T A G G C					h36b_a12.seq			
A T A T A G T A C T G C A A T A A A A T A A T A A T A C T C C C A A A T A T C A T A A A T G T T C C Majority									
	660	670	680	690	700				
651	A T A T A G T A C T G C A A T A A A A T A A T A A T A C T C C C A A A T A T C A T A A A T G T T C C					2603_a12.seq			
651	A T A T A G T A C T G C A A T A A A A T A A T A A T A C T C C C A A A T A T C A T A A A T G T T C C					18rs21_a12.seq			
651	A T A T A G T A C T G C A A T A A A A T A A T A A T A C T C C C A A A T A T C A T A A A T G T T C C					515_a12.seq			
651	A T A T A G T A C T G C A A T A A A A T A A T A A T A C T C C C A A A T A T C A T A A A T G T T C C					cjb111_a12.seq			
651	A T A T A G T A C T G C A A T A A A A T A A T A A T A C T C C C A A A T A T C A T A A A T G T T C C					h36b_a12.seq			
C A T C G A G T G C C C A C T G G G A A A C G A A T A G C C A C C T G C A A A T A C T A A A T G G G Majority									
	710	720	730	740	750				
701	C A T C G A G T G C C C A C T G G G A A A C G A A T A G C C A C C T G C A A A T A C T A A A T G G G					2603_a12.seq			
701	C A T C G A G T G C C C A C T G G G A A A C G A A T A G C C A C C T G C A A A T A C T A A A T G G G					18rs21_a12.seq			
701	C A T C G A G T G C C C A C T G G G A A A C G A A T A G C C A C C T G C A A A T A C T A A A T G G G					515_a12.seq			
701	C A T C G A G T G C C C A C T G G G A A A C G A A T A G C C A C C T G C A A A T A C T A A A T G G G					cjb111_a12.seq			
701	C A T C G A G T G C C C A C T G G G A A A C G A A T A G C C A C C T G C A A A T A C T A A A T G G G					h36b_a12.seq			
T T A A A G T T G G T C T T A C T C T T T G A A A A A T A A G T T T T A A A G A A A G T A T A C A T Majority									
	760	770	780	790	800				
751	T T A A A G T T G G T C T T A C T C T T T G A A A A A T A A G T T T T A A A G A A A G T A T A C A T					2603_a12.seq			
751	T T A A A G T T G G T C T T A C T C T T T G A A A A A T A A G T T T T A A A G A A A G T A T A C A T					18rs21_a12.seq			
751	T T A A A G T T G G T C T T A C T C T T T G A A A A A T A A G T T T T A A A G A A A G T A T A C A T					515_a12.seq			
751	T T A A A G T T G G T C T T A C T C T T T G A A A A A T A A G T T T T A A A G A A A G T A T A C A T					cjb111_a12.seq			
751	T T A A A G T T G G T C T T A C T C T T T G A A A A A T A A G T T T T A A A G A A A G T A T A C A T					h36b_a12.seq			
A T A C C A G A G A T A A T A G C A T T T A C T G C G A T A A A T C T A G C T T G A G G A T A C C A Majority									
	810	820	830	840	850				
801	A T A C C A G A G A T A A T A G C A T T T A C T G C G A T A A A T C T A G C T T G A G G A T A C C A					2603_a12.seq			
801	A T A C C A G A G A T A A T A G C A T T T A C T G C G A T A A A T C T A G C T T G A G G A T A C C A					18rs21_a12.seq			
801	A T A C C A G A G A T A A T A G C A T T T A C T G C G A T A A A T C T A G C T T G A G G A T A C C A					515_a12.seq			
801	A T A C C A G A G A T A A T A G C A T T T A C T G C G A T A A A T C T A G C T T G A G G A T A C C A					cjb111_a12.seq			
801	A T A C C A G A G A T A A T A G C A T T T A C T G C G A T A A A T C T A G C T T G A G G A T A C C A					h36b_a12.seq			
C T T C T T A A G C T A A C A G A A A G T G A C G C T C A T A A T C G C A A T A G C T A T C T G G C Majority									
	860	870	880	890	900				
851	C T T C T T A A G C T A A C A G A A A G T G A C G C T C A T A A T C G C A A T A G C T A T C T G G C					2603_a12.seq			
851	C T T C T T A A G C T A A C A G A A A G T G A C G C T C A T A A T C G C A A T A G C T A T C T G G C					18rs21_a12.seq			
851	C T T C T T A A G C T A A C A G A A A G T G A C G C T C A T A A T C G C A A T A G C T A T C T G G C					515_a12.seq			
851	C T T C T T A A G C T A A C A G A A A G T G A C G C T C A T A A T C G C A A T A G C T A T C T G G C					cjb111_a12.seq			
851	C T T C T T A A G C T A A C A G A A A G T G A C G C T C A T A A T C G C A A T A G C T A T C T G G C					h36b_a12.seq			

FIGURE 19A

58/487

TTACAGTATTACCAATCACAGTGAATTAACCTTGAAAAATCTTGTAGAAAGA Majority									
910 920 930 940 950									
901	TTACAGTATTACCAATCACAGTGAATTAACCTTGAAAAATCTTGTAGAAAGA 2603_a12.seq								
901	TTACAGTATTACCAATCACAGTGAATTAACCTTGAAAAATCTTGTAGAAAGA 18rs21_a12.seq								
901	TTACAGTATTACCAATCACAGTGAATTAACCTTGAAAAATCTTGTAGAAAGA 515_a12.seq								
901	TTACAGTATTACCAATCACAGTGAATTAACCTTGAAAAATCTTGTAGAAAGA cjb111_a12.seq								
901	TTACAGTATTACCAATCACAGTGAATTAACCTTGAAAAATCTTGTAGAAAGA h36b_a12.seq								
TTTGGCAACTGTCCTCTAACACTTTCTTGAATAGTTTGGTCAAATGAAAT Majority									
960 970 980 990 1000									
951	TTTGGCAACTGTCCTCTAACACTTTCTTGAATAGTTTGGTCAAATGAAAT 2603_a12.seq								
951	TTTGGCAACTGTCCTCTAACACTTTCTTGAATAGTTTGGTCAAATGAAAT 18rs21_a12.seq								
951	TTTGGCAACTGTCCTCTAACACTTTCTTGAATAGTTTGGTCAAATGAAAT 515_a12.seq								
951	TTTGGCAACTGTCCTCTAACACTTTCTTGAATAGTTTGGTCAAATGAAAT cjb111_a12.seq								
951	TTTGGCAACTGTCCTCTAACACTTTCTTGAATAGTTTGGTCAAATGAAAT h36b_a12.seq								
TACAGTGTGCGGGCCAATATTTGATGACCAATCCTAAACTGAAAAATAAGA Majority									
1010 1020 1030 1040 1050									
1001	TACAGTGTGCGGGCCAATATTTGATGACCAATCCTAAACTGAAAAATAAGA 2603_a12.seq								
1001	TACAGTGTGCGGGCCAATATTTGATGACCAATCCTAAACTGAAAAATAAGA 18rs21_a12.seq								
1001	TACAGTGTGCGGGCCAATATTTGATGACCAATCCTAAACTGAAAAATAAGA 515_a12.seq								
1001	TACAGTGTGCGGGCCAATATTTGATGACCAATCCTAAACTGAAAAATAAGA cjb111_a12.seq								
1001	TACAGTGTGCGGGCCAATATTTGATGACCAATCCTAAACTGAAAAATAAGA h36b_a12.seq								
TAATAGCAATAAATGCTTGAATAAGTTTACTATTTTGCAGAGATAACATT Majority									
1060 1070 1080 1090 1100									
1051	TAATAGCAATAAATGCTTGAATAAGTTTACTATTTTGCAGAGATAACATT 2603_a12.seq								
1051	TAATAGCAATAAATGCTTGAATAAGTTTACTATTTTGCAGAGATAACATT 18rs21_a12.seq								
1051	TAATAGCAATAAATGCTTGAATAAGTTTACTATTTTGCAGAGATAACATT 515_a12.seq								
1051	TAATAGCAATAAATGCTTGAATAAGTTTACTATTTTGCAGAGATAACATT cjb111_a12.seq								
1051	TAATAGCAATAAATGCTTGAATAAGTTTACTATTTTGCAGAGATAACATT h36b_a12.seq								
AGTCTTTTTTATATCTTTCTAATATTGGCAAACAAGCCACGTAAGTTAGAT Majority									
1110 1120 1130 1140 1150									
1101	AGTCTTTTTTATATCTTTCTAATATTGGCAAACAAGCCACGTAAGTTAGAT 2603_a12.seq								
1101	AGTCTTTTTTATATCTTTCTAATATTGGCAAACAAGCCACGTAAGTTAGAT 18rs21_a12.seq								
1101	AGTCTTTTTTATATCTTTCTAATATTGGCAAACAAGCCACGTAAGTTAGAT 515_a12.seq								
1101	AGTCTTTTTTATATCTTTCTAATATTGGCAAACAAGCCACGTAAGTTAGAT cjb111_a12.seq								
1101	AGTCTTTTTTATATCTTTCTAATATTGGCAAACAAGCCACGTAAGTTAGAT h36b_a12.seq								
AGAAAAACAATCGAAATTAAATTCCTCAACGATATTAAATGGAATAACC Majority									
1160 1170 1180 1190 1200									
1151	AGAAAAACAATCGAAATTAAATTCCTCAACGATATTAAATGGAATAACC 2603_a12.seq								
1151	AGAAAAACAATCGAAATTAAATTCCTCAACGATATTAAATGGAATAACC 18rs21_a12.seq								
1151	AGAAAAACAATCGAAATTAAATTCCTCAACGATATTAAATGGAATAACC 515_a12.seq								
1151	AGAAAAACAATCGAAATTAAATTCCTCAACGATATTAAATGGAATAACC cjb111_a12.seq								
1151	AGAAAAACAATCGAAATTAAATTCCTCAACGATATTAAATGGAATAACC h36b_a12.seq								
ATTGTTTAAAAGGTAATTGCTTACACCAATAAATGTTCTGATATCAAAGTT Majority									
1210 1220 1230 1240 1250									
1201	ATTGTTTAAAAGGTAATTGCTTACACCAATAAATGTTCTGATATCAAAGTT 2603_a12.seq								
1201	ATTGTTTAAAAGGTAATTGCTTACACCAATAAATGTTCTGATATCAAAGTT 18rs21_a12.seq								
1201	ATTGTTTAAAAGGTAATTGCTTACACCAATAAATGTTCTGATATCAAAGTT 515_a12.seq								
1201	ATTGTTTAAAAGGTAATTGCTTACACCAATAAATGTTCTGATATCAAAGTT cjb111_a12.seq								
1201	ATTGTTTAAAAGGTAATTGCTTACACCAATAAATGTTCTGATATCAAAGTT h36b_a12.seq								
AGCAAATATATAGCATACAAAGGAATCGCAAAGACATAGTTGAGAGCTACCA Majority									
1260 1270 1280 1290 1300									
1251	AGCAAATATATAGCATACAAAGGAATCGCAAAGACATAGTTGAGAGCTACCA 2603_a12.seq								
1251	AGCAAATATATAGCATACAAAGGAATCGCAAAGACATAGTTGAGAGCTACCA 18rs21_a12.seq								
1251	AGCAAATATATAGCATACAAAGGAATCGCAAAGACATAGTTGAGAGCTACCA 515_a12.seq								
1251	AGCAAATATATAGCATACAAAGGAATCGCAAAGACATAGTTGAGAGCTACCA cjb111_a12.seq								
1251	AGCAAATATATAGCATACAAAGGAATCGCAAAGACATAGTTGAGAGCTACCA h36b_a12.seq								
TAGATACGGTCAAGCTAACTGTACCAAATAGACTAGCTTTAATAAAATCT Majority									
1310 1320 1330 1340 1350									
1301	TAGATACGGTCAAGCTAACTGTACCAAATAGACTAGCTTTAATAAAATCT 2603_a12.seq								
1301	TAGATACGGTCAAGCTAACTGTACCAAATAGACTAGCTTTAATAAAATCT 18rs21_a12.seq								
1301	TAGATACGGTCAAGCTAACTGTACCAAATAGACTAGCTTTAATAAAATCT 515_a12.seq								
1301	TAGATACGGTCAAGCTAACTGTACCAAATAGACTAGCTTTAATAAAATCT cjb111_a12.seq								
1301	TAGATACGGTCAAGCTAACTGTACCAAATAGACTAGCTTTAATAAAATCT h36b_a12.seq								

FIGURE 19B

59/487

TTTGCACCTCTCTCTATTTTTTCCAGAAAAATAGCGAAACTTGCTAAAAATAA Majority									
1360 1370 1380 1390 1400									
1351	TTTGCACCTCTCTCTATTTTTTCCAGAAAAATAGCGAAACTTGCTAAAAATAA	2603_a12.seq							
1351	TTTGCACCTCTCTCTATTTTTTCCAGAAAAATAGCGAAACTTGCTAAAAATAA	18rs21_a12.seq							
1351	TTTGCACCTCTCTCTATTTTTTCCAGAAAAATAGCGAAACTTGCTAAAAATAA	515_a12.seq							
1351	TTTGCACCTCTCTCTATTTTTTCCAGAAAAATAGCGAAACTTGCTAAAAATAA	cjb111_a12.seq							
1351	TTTGCACCTCTCTCTATTTTTTCCAGAAAAATAGCGAAACTTGCTAAAAATAA	h36b_a12.seq							
AGCTAGAGCAACCATATTCATCGGTAAACCAATAAACGTTTCTGGACCAC Majority									
1410 1420 1430 1440 1450									
1401	AGCTAGAGCAACCATATTCATCGGTAAACCAATAAACGTTTCTGGACCAC	2603_a12.seq							
1401	AGCTAGAGCAACCATATTCATCGGTAAACCAATAAACGTTTCTGGACCAC	18rs21_a12.seq							
1401	AGCTAGAGCAACCATATTCATCGGTAAACCAATAAACGTTTCTGGACCAC	515_a12.seq							
1401	AGCTAGAGCAACCATATTCATCGGTAAACCAATAAACGTTTCTGGACCAC	cjb111_a12.seq							
1401	AGCTAGAGCAACCATATTCATCGGTAAACCAATAAACGTTTCTGGACCAC	h36b_a12.seq							
GATTAGCAAGTATAACTTTTTAAAGTGATCTTAATAAGAGTACACCATAA Majority									
1460 1470 1480 1490 1500									
1451	GATTAGCAAGTATAACTTTTTAAAGTGATCTTAATAAGAGTACACCATAA	2603_a12.seq							
1451	GATTAGCAAGTATAACTTTTTAAAGTGATCTTAATAAGAGTACACCATAA	18rs21_a12.seq							
1451	GATTAGCAAGTATAACTTTTTAAAGTGATCTTAATAAGAGTACACCATAA	515_a12.seq							
1451	GATTAGCAAGTATAACTTTTTAAAGTGATCTTAATAAGAGTACACCATAA	cjb111_a12.seq							
1451	GATTAGCAAGTATAACTTTTTAAAGTGATCTTAATAAGAGTACACCATAA	h36b_a12.seq							
CTTGATTTTCAAATCAAATAAAATAAAAGCAACTAACATCGGAAGGATTGA Majority									
1510 1520 1530 1540 1550									
1501	CTTGATTTTCAAATCAAATAAAATAAAAGCAACTAACATCGGAAGGATTGA	2603_a12.seq							
1501	CTTGATTTTCAAATCAAATAAAATAAAAGCAACTAACATCGGAAGGATTGA	18rs21_a12.seq							
1501	CTTGATTTTCAAATCAAATAAAATAAAAGCAACTAACATCGGAAGGATTGA	515_a12.seq							
1501	CTTGATTTTCAAATCAAATAAAATAAAAGCAACTAACATCGGAAGGATTGA	cjb111_a12.seq							
1501	CTTGATTTTCAAATCAAATAAAATAAAAGCAACTAACATCGGAAGGATTGA	h36b_a12.seq							
AAAATCAACCTTTTAAAAAATTCTGCTCCTGGTATTAATGGAAATGAAACCA Majority									
1560 1570 1580 1590 1600									
1551	AAAATCAACCTTTTAAAAAATTCTGCTCCTGGTATTAATGGAAATGAAACCA	2603_a12.seq							
1551	AAAATCAACCTTTTAAAAAATTCTGCTCCTGGTATTAATGGAAATGAAACCA	18rs21_a12.seq							
1551	AAAATCAACCTTTTAAAAAATTCTGCTCCTGGTATTAATGGAAATGAAACCA	515_a12.seq							
1551	AAAATCAACCTTTTAAAAAATTCTGCTCCTGGTATTAATGGAAATGAAACCA	cjb111_a12.seq							
1551	AAAATCAACCTTTTAAAAAATTCTGCTCCTGGTATTAATGGAAATGAAACCA	h36b_a12.seq							
TCATCAATACAAAAGATAAGGCCAGAAAGAATGGCGATTGTACACCATTTTA Majority									
1610 1620 1630 1640 1650									
1601	TCATCAATACAAAAGATAAGGCCAGAAAGAATGGCGATTGTACACCATTTTA	2603_a12.seq							
1601	TCATCAATACAAAAGATAAGGCCAGAAAGAATGGCGATTGTACACCATTTTA	18rs21_a12.seq							
1601	TCATCAATACAAAAGATAAGGCCAGAAAGAATGGCGATTGTACACCATTTTA	515_a12.seq							
1601	TCATCAATACAAAAGATAAGGCCAGAAAGAATGGCGATTGTACACCATTTTA	cjb111_a12.seq							
1601	TCATCAATACAAAAGATAAGGCCAGAAAGAATGGCGATTGTACACCATTTTA	h36b_a12.seq							
CGTGTATTTGTCTATAAAAAAATTCTCCCAATTTTAAATAAATTGAAAGAAAG Majority									
1660 1670 1680 1690 1700									
1651	CGTGTATTTGTCTATAAAAAAATTCTCCCAATTTTAAATAAATTGAAAGAAAG	2603_a12.seq							
1651	CGTGTATTTGTCTATAAAAAAATTCTCCCAATTTTAAATAAATTGAAAGAAAG	18rs21_a12.seq							
1651	CGTGTATTTGTCTATAAAAAAATTCTCCCAATTTTAAATAAATTGAAAGAAAG	515_a12.seq							
1651	CGTGTATTTGTCTATAAAAAAATTCTCCCAATTTTAAATAAATTGAAAGAAAG	cjb111_a12.seq							
1651	CGTGTATTTGTCTATAAAAAAATTCTCCCAATTTTAAATAAATTGAAAGAAAG	h36b_a12.seq							
CTCCAAAGGTAAGCGTATGTACGCCGAAAAAA - CCTTTGTCTTCTCCCATC Majority									
1710 1720 1730 1740 1750									
1701	CTCCAAAGGTAAGCGTATGTACGCCGAAAAAA - CCTTTGTCTTCTCCCATC	2603_a12.seq							
1701	CTCCAAAGGTAAGCGTATGTACGCCGAAAAAA - CCTTTGTCTTCTCCCATC	18rs21_a12.seq							
1701	CTCCAAAGGTAAGCGTATGTACGCCGAAAAAA - CCTTTGTCTTCTCCCATC	515_a12.seq							
1701	CTCCAAAGGTAAGCGTATGTACGCCGAAAAAA - CCTTTGTCTTCTCCCATC	cjb111_a12.seq							
1701	CTCCAAAGGTAAGCGTATGTACGCCGAAAAAA - CCTTTGTCTTCTCCCATC	h36b_a12.seq							
CAGACTTTTACTGTCGGTTGTGGAATCTCACACATCAGCTTTTCGCTCGCG Majority									
1760 1770 1780 1790 1800									
1750	CAGACTTTTACTGTCGGTTGTGGAATCTCACACATCAGCTTTTCGCTCGCG	2603_a12.seq							
1750	CAGACTTTTACTGTCGGTTGTGGAATCTCACACATCAGCTTTTCGCTCGCG	18rs21_a12.seq							
1751	CAGACTTTTACTGTCGGTTGTGGAATCTCACACATCAGCTTTTCGCTCGCG	515_a12.seq							
1750	CAGACTTTTACTGTCGGTTGTGGAATCTCACACATCAGCTTTTCGCTCGCG	cjb111_a12.seq							
1750	CAGACTTTTACTGTCGGTTGTGGAATCTCACACATCAGCTTTTCGCTCGCG	h36b_a12.seq							

FIGURE 19C

G A C T G A T G C T T C A C A A C T G A C A A A T A A G T T G G A A G C G A T T A C C G C C G G T C Majority									
1810 1820 1830 1840 1850									
1800	G A C T G A T G C T T C A C A A C T G A C A A A T A A G T T G G A A G C G A T T A C C G C C G G T C 2603_a12.seq								
1800	G A C T G A T G C T T C A C A A C T G A C A A A T A A G T T G G A A G C G A T T A C C G C C G G T C 18rs21_a12.seq								
1801	G A C T G A T G C T T C A C A A C T G A C A A A T A A G T T G G A A G C G A T T A C C G C C G G T C 515_a12.seq								
1800	G A C T G A T G C T T C A C A A C T G A C A A A T A A G T T G G A A G C G A T T A C C G C C G G T C cjb111_a12.seq								
1800	G A C T G A T G C T T C A C A A C T G A C A A A T A A G T T G G A A G C G A T T A C C G C C G G T C h36b_a12.seq								
G G G A A T T A C A C C C T G C C C T G A A G A C A C C T A T A G C A T A A C A A A A A A A C T T Majority									
1860 1870 1880 1890 1900									
1850	G G G A A T T A C A C C C T G C C C T G A A G A C A C C T A T A G C A T A A C A A A A A A A C T T 2603_a12.seq								
1850	G G G A A T T A C A C C C T G C C C T G A A G A C A C C T A T A G C A T A A C A A A A A A A C T T 18rs21_a12.seq								
1851	G G G A A T T A C A C C C T G C C C T G A A G A C A C C T A T A G C A T A A C A A A A A A A C T T 515_a12.seq								
1850	G G G A A T T A C A C C C T G C C C T G A A G A C A C C T A T A G C A T A A C A A A A A A A C T T cjb111_a12.seq								
1850	G G G A A T T A C A C C C T G C C C T G A A G A C A C C T A T A G C A T A A C A A A A A A A C T T h36b_a12.seq								
G C A A T T G C A A G T T T T T T A A T T A C T A A T T A G T A G T A G T G A T T A A A A A T C A T Majority									
1910 1920 1930 1940 1950									
1900	G C A A T T G C A A G T T T T T T A A T T A C T A A T T A G T A G T A G T G A T T A A A A A T C A T 2603_a12.seq								
1900	G C A A T T G C A A G T T T T T T A A T T A C T A A T T A G T A G T A G T G A T T A A A A A T C A T 18rs21_a12.seq								
1901	G C A A T T G C A A G T T T T T T A A T T A C T A A T T A G T A G T A G T G A T T A A A A A T C A T 515_a12.seq								
1900	G C A A T T G C A A G T T T T T T A A T T A C T A A T T A G T A G T A G T G A T T A A A A A T C A T cjb111_a12.seq								
1900	G C A A T T G C A A G T T T T T T A A T T A C T A A T T A G T A G T A G T G A T T A A A A A T C A T h36b_a12.seq								
A T T A A T A C C A A A T T A C T A T A C T G T A T C G T T T C T T T C A G A T T T G C T A T T T T Majority									
1960 1970 1980 1990 2000									
1950	A T T A A T A C C A A A T T A C T A T A C T G T A T C G T T T C T T T C A G A T T T G C T A T T T T 2603_a12.seq								
1950	A T T A A T A C C A A A T T A C T A T A C T G T A T C G T T T C T T T C A G A T T T G C T A T T T T 18rs21_a12.seq								
1951	A T T A A T A C C A A A T T A C T A T A C T G T A T C G T T T C T T T C A G A T T T G C T A T T T T 515_a12.seq								
1950	A T T A A T A C C A A A T T A C T A T A C T G T A T C G T T T C T T T C A G A T T T G C T A T T T T cjb111_a12.seq								
1950	A T T A A T A C C A A A T T A C T A T A C T G T A T C G T T T C T T T C A G A T T T G C T A T T T T h36b_a12.seq								
T A G T T T T T C T T A A A A A G A T A A A C A A A A T T C C C A A A A T A A T A C A A C C A A G A Majority									
2010 2020 2030 2040 2050									
2000	T A G T T T T T C T T A A A A A G A T A A A C A A A A T T C C C A A A A T A A T A C A A C C A A G A 2603_a12.seq								
2000	T A G T T T T T C T T A A A A A G A T A A A C A A A A T T C C C A A A A T A A T A C A A C C A A G A 18rs21_a12.seq								
2001	T A G T T T T T C T T A A A A A G A T A A A C A A A A T T C C C A A A A T A A T A C A A C C A A G A 515_a12.seq								
2000	T A G T T T T T C T T A A A A A G A T A A A C A A A A T T C C C A A A A T A A T A C A A C C A A G A cjb111_a12.seq								
2000	T A G T T T T T C T T A A A A A G A T A A A C A A A A T T C C C A A A A T A A T A C A A C C A A G A h36b_a12.seq								
A T T G T C A G T C C T C C A C C A A T A A T C A T T C C T G T T T T A G G A A G A A A T G A T T C Majority									
2060 2070 2080 2090 2100									
2050	A T T G T C A G T C C T C C A C C A A T A A T C A T T C C T G T T T T A G G A A G A A A T G A T T C 2603_a12.seq								
2050	A T T G T C A G T C C T C C A C C A A T A A T C A T T C C T G T T T T A G G A A G A A A T G A T T C 18rs21_a12.seq								
2051	A T T G T C A G T C C T C C A C C A A T A A T C A T T C C T G T T T T A G G A A G A A A T G A T T C 515_a12.seq								
2050	A T T G T C A G T C C T C C A C C A A T A A T C A T T C C T G T T T T A G G A A G A A A T G A T T C cjb111_a12.seq								
2050	A T T G T C A G T C C T C C A C C A A T A A T C A T T C C T G T T T T A G G A A G A A A T G A T T C h36b_a12.seq								
T G G A A A A A G C G G T T G T G A T G C T T T A G G A T T T G T T C G T G C A G G A G T T T C T T Majority									
2110 2120 2130 2140 2150									
2100	T G G A A A A A G C G G T T G T G A T G C T T T A G G A T T T G T T C G T G C A G G A G T T T C T T 2603_a12.seq								
2100	T G G A A A A A G C G G T T G T G A T G C T T T A G G A T T T G T T C G T G C A G G A G T T T C T T 18rs21_a12.seq								
2101	T G G A A A A A G C G G T T G T G A T G C T T T A G G A T T T G T T C G T G C A G G A G T T T C T T 515_a12.seq								
2100	T G G A A A A A G C G G T T G T G A T G C T T T A G G A T T T G T T C G T G C A G G A G T T T C T T cjb111_a12.seq								
2100	T G G A A A A A G C G G T T G T G A T G C T T T A G G A T T T G T T C G T G C A G G A G T T T C T T h36b_a12.seq								
T T T C G T T T T C T A C C T C T A C T T C C T G T G T T T T A T T A G C A A C T A C A G C A A C T Majority									
2160 2170 2180 2190 2200									
2150	T T T C G T T T T C T A C C T C T A C T T C C T G T G T T T T A T T A G C A A C T A C A G C A A C T 2603_a12.seq								
2150	T T T C G T T T T C T A C C T C T A C T T C C T G T G T T T T A T T A G C A A C T A C A G C A A C T 18rs21_a12.seq								
2151	T T T C G T T T T C T A C C T C T A C T T C C T G T G T T T T A T T A G C A A C T A C A G C A A C T 515_a12.seq								
2150	T T T C G T T T T C T A C C T C T A C T T C C T G T G T T T T A T T A G C A A C T A C A G C A A C T cjb111_a12.seq								
2150	T T T C G T T T T C T A C C T C T A C T T C C T G T G T T T T A T T A G C A A C T A C A G C A A C T h36b_a12.seq								
A C A G C A T C C T T C A T A G A T A T A C G G T A A C C A G T T A G T G C T T T T G C T T C T C G Majority									
2210 2220 2230 2240 2250									
2200	A C A G C A T C C T T C A T A G A T A T A C G G T A A C C A G T T A G T G C T T T T G C T T C T C G 2603_a12.seq								
2200	A C A G C A T C C T T C A T A G A T A T A C G G T A A C C A G T T A G T G C T T T T G C T T C T C G 18rs21_a12.seq								
2201	A C A G C A T C C T T C A T A G A T A T A C G G T A A C C A G T T A G T G C T T T T G C T T C T C G 515_a12.seq								
2200	A C A G C A T C C T T C A T A G A T A T A C G G T A A C C A G T T A G T G C T T T T G C T T C T C G cjb111_a12.seq								
2200	A C A G C A T C C T T C A T A G A T A T A C G G T A A C C A G T T A G T G C T T T T G C T T C T C G h36b_a12.seq								

FIGURE 19D

61/487

	AAAAAATAAGTAACTAATTAACCTTCAACCTCAATTTCTCCCTTAT	Majority
	2260 2270 2280 2290 2300	
2250	AAAAATATACCTTACCAGGTAATAAACCTTCAACCTCAATTTCTCCCTTAT	2603_a12.seq
2250	AAAAATATACCTTACCAGGTAATAAACCTTCAACCTCAATTTCTCCCTTAT	18rs21_a12.seq
2251	AAAAATATACCTTACCAGGTAATAAACCTTCAACCTCAATTTCTCCCTTAT	515_a12.seq
2250	AAAAATATACCTTACCAGGTAATAAACCTTCAACCTCAATTTCTCCCTTAT	cjb111_a12.seq
2250	AAAAATATACCTTACCAGGTAATAAACCTTCAACCTCAATTTCTCCCTTAT	h36b_a12.seq
	CATCAGTTACTAATGAAGTAATCCCATCTTGATCGGTCGTAAATCGTCCA	Majority
	2310 2320 2330 2340 2350	
2300	CATCAGTTACTAATGAAGTAATCCCATCTTGATCGGTCGTAAATCGTCCA	2603_a12.seq
2300	CATCAGTTACTAATGAAGTAATCCCATCTTGATCGGTCGTAAATCGTCCA	18rs21_a12.seq
2301	CATCAGTTACTAATGAAGTAATCCCATCTTGATCGGTCGTAAATCGTCCA	515_a12.seq
2300	CATCAGTTACTAATGAAGTAATCCCATCTTGATCGGTCGTAAATCGTCCA	cjb111_a12.seq
2300	CATCAGTTACTAATGAAGTAATCCCATCTTGATCGGTCGTAAATCGTCCA	h36b_a12.seq
	TTTTTAAAGCGAAGCTGGCTGATTCTCGTTATCGTATAAATACAAATATTAC	Majority
	2360 2370 2380 2390 2400	
2350	TTTTTAAAGCGAAGCTGGCTGATTCTCGTTATCGTATAAATACAAATATTAC	2603_a12.seq
2350	TTTTTAAAGCGAAGCTGGCTGATTCTCGTTATCGTATAAATACAAATATTAC	18rs21_a12.seq
2351	TTTTTAAAGCGAAGCTGGCTGATTCTCGTTATCGTATAAATACAAATATTAC	515_a12.seq
2350	TTTTTAAAGCGAAGCTGGCTGATTCTCGTTATCGTATAAATACAAATATTAC	cjb111_a12.seq
2350	TTTTTAAAGCGAAGCTGGCTGATTCTCGTTATCGTATAAATACAAATATTAC	h36b_a12.seq
	TCCGGATAGCCCTTTTCTTTATCTTTCTTCTTTTGTATATTTAATAAGTT	Majority
	2410 2420 2430 2440 2450	
2400	TCCGGATAGCCCTTTTCTTTATCTTTCTTCTTTTGTATATTTAATAAGTT	2603_a12.seq
2400	TCCGGATAGCCCTTTTCTTTATCTTTCTTCTTTTGTATATTTAATAAGTT	18rs21_a12.seq
2401	TCCGGATAGCCCTTTTCTTTATCTTTCTTCTTTTGTATATTTAATAAGTT	515_a12.seq
2400	TCCGGATAGCCCTTTTCTTTATCTTTCTTCTTTTGTATATTTAATAAGTT	cjb111_a12.seq
2400	TCCGGATAGCCCTTTTCTTTATCTTTCTTCTTTTGTATATTTAATAAGTT	h36b_a12.seq
	TAAATCGGCCCTGTTTCAACTTTTCGCTTAGGATTTATCTGTAATTGATTT	Majority
	2460 2470 2480 2490 2500	
2450	TAAATCGGCCCTGTTTCAACTTTTCGCTTAGGATTTATCTGTAATTGATTT	2603_a12.seq
2450	TAAATCGGCCCTGTTTCAACTTTTCGCTTAGGATTTATCTGTAATTGATTT	18rs21_a12.seq
2451	TAAATCGGCCCTGTTTCAACTTTTCGCTTAGGATTTATCTGTAATTGATTT	515_a12.seq
2450	TAAATCGGCCCTGTTTCAACTTTTCGCTTAGGATTTATCTGTAATTGATTT	cjb111_a12.seq
2450	TAAATCGGCCCTGTTTCAACTTTTCGCTTAGGATTTATCTGTAATTGATTT	h36b_a12.seq
	GATAACTTATCATCTGGTAATTCAATATAAAAGGTACTATTGTTGAAAC	Majority
	2510 2520 2530 2540 2550	
2500	GATAACTTATCATCTGGTAATTCAATATAAAAGGTACTATTGTTGAAAC	2603_a12.seq
2500	GATAACTTATCATCTGGTAATTCAATATAAAAGGTACTATTGTTGAAAC	18rs21_a12.seq
2501	GATAACTTATCATCTGGTAATTCAATATAAAAGGTACTATTGTTGAAAC	515_a12.seq
2500	GATAACTTATCATCTGGTAATTCAATATAAAAGGTACTATTGTTGAAAC	cjb111_a12.seq
2500	GATAACTTATCATCTGGTAATTCAATATAAAAGGTACTATTGTTGAAAC	h36b_a12.seq
	GCTTTGATCAGCTTTTATAAGCACGACCAAAGTACGAACCATTTGGGAGTG	Majority
	2560 2570 2580 2590 2600	
2550	GCTTTGATCAGCTTTTATAAGCACGACCAAAGTACGAACCATTTGGGAGTG	2603_a12.seq
2550	GCTTTGATCAGCTTTTATAAGCACGACCAAAGTACGAACCATTTGGGAGTG	18rs21_a12.seq
2551	GCTTTGATCAGCTTTTATAAGCACGACCAAAGTACGAACCATTTGGGAGTG	515_a12.seq
2550	GCTTTGATCAGCTTTTATAAGCACGACCAAAGTACGAACCATTTGGGAGTG	cjb111_a12.seq
2550	GCTTTGATCAGCTTTTATAAGCACGACCAAAGTACGAACCATTTGGGAGTG	h36b_a12.seq
	CTATCTTTGTCTGACCAATTAGTATCAGTAGGAGAAAGTCAAGATACTCTTA	Majority
	2610 2620 2630 2640 2650	
2600	CTATCTTTGTCTGACCAATTAGTATCAGTAGGAGAAAGTCAAGATACTCTTA	2603_a12.seq
2600	CTATCTTTGTCTGACCAATTAGTATCAGTAGGAGAAAGTCAAGATACTCTTA	18rs21_a12.seq
2601	CTATCTTTGTCTGACCAATTAGTATCAGTAGGAGAAAGTCAAGATACTCTTA	515_a12.seq
2600	CTATCTTTGTCTGACCAATTAGTATCAGTAGGAGAAAGTCAAGATACTCTTA	cjb111_a12.seq
2600	CTATCTTTGTCTGACCAATTAGTATCAGTAGGAGAAAGTCAAGATACTCTTA	h36b_a12.seq
	TACTTCTGGTTCAATTTCGCTATCTGTCAATTTGGCTCAATAAATCAACTTT	Majority
	2660 2670 2680 2690 2700	
2650	TACTTCTGGTTCAATTTCGCTATCTGTCAATTTGGCTCAATAAATCAACTTT	2603_a12.seq
2650	TACTTCTGGTTCAATTTCGCTATCTGTCAATTTGGCTCAATAAATCAACTTT	18rs21_a12.seq
2651	TACTTCTGGTTCAATTTCGCTATCTGTCAATTTGGCTCAATAAATCAACTTT	515_a12.seq
2650	TACTTCTGGTTCAATTTCGCTATCTGTCAATTTGGCTCAATAAATCAACTTT	cjb111_a12.seq
2650	TACTTCTGGTTCAATTTCGCTATCTGTCAATTTGGCTCAATAAATCAACTTT	h36b_a12.seq

FIGURE 19E

62/487

	T A A G T T G T C A G T C A C A G T C C A T A A A C G A T A A G A A A T C C C C T C C T C T G T A G Majority									
	2710		2720		2730		2740		2750	
2700	T A A G T T G T C A G T C A C A G T C C A T A A A C G A T A A G A A A T C C C C T C C T C T G T A G 2603_a12.seq									
2700	T A A G T T G T C A G T C A C A G T C C A T A A A C G A T A A G A A A T C C C C T C C T C T G T A G 18rs21_a12.seq									
2701	T A A G T T G T C A G T C A C A G T C C A T A A A C G A T A A G A A A T C C C C T C C T C T G T A G 515_a12.seq									
2700	T A A G T T G T C A G T C A C A G T C C A T A A A C G A T A A G A A A T C C C C T C C T C T G T A G cjb111_a12.seq									
2700	T A A G T T G T C A G T C A C A G T C C A T A A A C G A T A A G A A A T C C C C T C C T C T G T A G h36b_a12.seq									
	T A T T T G G C T G A A G T C C T A T C T G T G T G A T T G T T A G T T G A T T A G G G G T A T C A Majority									
	2760		2770		2780		2790		2800	
2750	T A T T T G G C T G A A G T C C T A T C T G T G T G A T T G T T A G T T G A T T A G G G G T A T C A 2603_a12.seq									
2750	T A T T T G G C T G A A G T C C T A T C T G T G T G A T T G T T A G T T G A T T A G G G G T A T C A 18rs21_a12.seq									
2751	T A T T T G G C T G A A G T C C T A T C T G T G T G A T T G T T A G T T G A T T A G G G G T A T C A 515_a12.seq									
2750	T A T T T G G C T G A A G T C C T A T C T G T G T G A T T G T T A G T T G A T T A G G G G T A T C A cjb111_a12.seq									
2750	T A T T T G G C T G A A G T C C T A T C T G T G T G A T T G T T A G T T G A T T A G G G G T A T C A h36b_a12.seq									
	G C A T T T T A C A C T G G C T A C C G A A A A A A A C G C T A A T T G T A C C A A T C C T A A A A A Majority									
	2810		2820		2830		2840		2850	
2800	G C A T T T T A C A C T G G C T A C C G A A A A A A A C G C T A A T T G T A C C A A T C C T A A A A A 2603_a12.seq									
2800	G C A T T T T A C A C T G G C T A C C G A A A A A A A C G C T A A T T G T A C C A A T C C T A A A A A 18rs21_a12.seq									
2801	G C A T T T T A C A C T G G C T A C C G A A A A A A A C G C T A A T T G T A C C A A T C C T A A A A A 515_a12.seq									
2800	G C A T T T T A C A C T G G C T A C C G A A A A A A A C G C T A A T T G T A C C A A T C C T A A A A A cjb111_a12.seq									
2800	G C A T T T T A C A C T G G C T A C C G A A A A A A A C G C T A A T T G T A C C A A T C C T A A A A A h36b_a12.seq									
	G C A A C A T A G T A G A A G T C C T A A A C T T T T T C T A A T C T T T T T C A T T T T T G A T T Majority									
	2860		2870		2880		2890		2900	
2850	G C A A C A T A G T A G A A G T C C T A A A C T T T T T C T A A T C T T T T T C A T T T T T G A T T 2603_a12.seq									
2850	G C A A C A T A G T A G A A G T C C T A A A C T T T T T C T A A T C T T T T T C A T T T T T G A T T 18rs21_a12.seq									
2851	G C A A C A T A G T A G A A G T C C T A A A C T T T T T C T A A T C T T T T T C A T T T T T G A T T 515_a12.seq									
2850	G C A A C A T A G T A G A A G T C C T A A A C T T T T T C T A A T C T T T T T C A T T T T T G A T T cjb111_a12.seq									
2850	G C A A C A T A G T A G A A G T C C T A A A C T T T T T C T A A T C T T T T T C A T T T T T G A T T h36b_a12.seq									
	T C C C T T T C T C T T T C T C T T T T A A A T T T T C G T T T T A A A T A T A A T A G T A A A G Majority									
	2910		2920		2930		2940		2950	
2900	T C C C T T T C T C T T T C T C T T T T A A A T T T T C G T T T T A A A T A T A A T A G T A A A G 2603_a12.seq									
2900	T C C C T T T C T C T T T C T C T T T T A A A T T T T C G T T T T A A A T A T A A T A G T A A A G 18rs21_a12.seq									
2901	T C C C T T T C T C T T T C T C T T T T A A A T T T T C G T T T T A A A T A T A A T A G T A A A G 515_a12.seq									
2900	T C C C T T T C T C T T T C T C T T T T A A A T T T T C G T T T T A A A T A T A A T A G T A A A G cjb111_a12.seq									
2900	T C C C T T T C T C T T T C T C T T T T A A A T T T T C G T T T T A A A T A T A A T A G T A A A G h36b_a12.seq									
	C G A C T A A T A T A A G A A T A A C T A G G A T T G A T A A G A G G A A A T A A A G T T T A T A G Majority									
	2960		2970		2980		2990		3000	
2950	C G A C T A A T A T A A G A A T A A C T A G G A T T G A T A A G A G G A A A T A A A G T T T A T A G 2603_a12.seq									
2950	C G A C T A A T A T A A G A A T A A C T A G G A T T G A T A A G A G G A A A T A A A G T T T A T A G 18rs21_a12.seq									
2951	C G A C T A A T A T A A G A A T A A C T A G G A T T G A T A A G A G G A A A T A A A G T T T A T A G 515_a12.seq									
2950	C G A C T A A T A T A A G A A T A A C T A G G A T T G A T A A G A G G A A A T A A A G T T T A T A G cjb111_a12.seq									
2950	C G A C T A A T A T A A G A A T A A C T A G G A T T G A T A A G A G G A A A T A A A G T T T A T A G h36b_a12.seq									
	T G T G T T T G C A A T T C T T T C A T T A A A T A G T T C T T T T C T T T A A C A G G A G G T A C Majority									
	3010		3020		3030		3040		3050	
3000	T G T G T T T G C A A T T C T T T C A T T A A A T A G T T C T T T T C T T T A A C A G G A G G T A C 2603_a12.seq									
3000	T G T G T T T G C A A T T C T T T C A T T A A A T A G T T C T T T T C T T T A A C A G G A G G T A C 18rs21_a12.seq									
3001	T G T G T T T G C A A T T C T T T C A T T A A A T A G T T C T T T T C T T T A A C A G G A G G T A C 515_a12.seq									
3000	T G T G T T T G C A A T T C T T T C A T T A A A T A G T T C T T T T C T T T A A C A G G A G G T A C cjb111_a12.seq									
3000	T G T G T T T G C A A T T C T T T C A T T A A A T A G T T C T T T T C T T T A A C A G G A G G T A C h36b_a12.seq									
	A T A C T T G A T T C G A T G C C C T C T A A C T A G T A A A C G A T G T G A A T T A A T C G A A T Majority									
	3060		3070		3080		3090		3100	
3050	A T A C T T G A T T C G A T G C C C T C T A A C T A G T A A A C G A T G T G A A T T A A T C G A A T 2603_a12.seq									
3050	A T A C T T G A T T C G A T G C C C T C T A A C T A G T A A A C G A T G T G A A T T A A T C G A A T 18rs21_a12.seq									
3051	A T A C T T G A T T C G A T G C C C T C T A A C T A G T A A A C G A T G T G A A T T A A T C G A A T 515_a12.seq									
3050	A T A C T T G A T T C G A T G C C C T C T A A C T A G T A A A C G A T G T G A A T T A A T C G A A T cjb111_a12.seq									
3050	A T A C T T G A T T C G A T G C C C T C T A A C T A G T A A A C G A T G T G A A T T A A T C G A A T h36b_a12.seq									
	A A G G T G T A C A T G T T A G C A A A G T C G C A T A A T C C T T A C C T T T A A C A A C C A A T Majority									
	3110		3120		3130		3140		3150	
3100	A A G G T G T A C A T G T T A G C A A A G T C G C A T A A T C C T T A C C T T T A A C A A C C A A T 2603_a12.seq									
3100	A A G G T G T A C A T G T T A G C A A A G T C G C A T A A T C C T T A C C T T T A A C A A C C A A T 18rs21_a12.seq									
3101	A A G G T G T A C A T G T T A G C A A A G T C G C A T A A T C C T T A C C T T T A A C A A C C A A T 515_a12.seq									
3100	A A G G T G T A C A T G T T A G C A A A G T C G C A T A A T C C T T A C C T T T A A C A A C C A A T cjb111_a12.seq									
3100	A A G G T G T A C A T G T T A G C A A A G T C G C A T A A T C C T T A C C T T T A A C A A C C A A T h36b_a12.seq									

FIGURE 19F

63/487

		A A T T T A G A A A A A T T A T C T G G C T T T A C A A C A C T T A T T T G A T C A A C C T T A T A Majority									
		3160		3170		3180		3190		3200	
3150	A A T T T A G A A A A A T T A T C T G G C T T T A C A A C A C T T A T T T G A T C A A C C T T A T A	2603_al2.seq									
3150	A A T T T A G A A A A A T T A T C T G G C T T T A C A A C A C T T A T T T G A T C A A C C T T A T A	18rs21_al2.seq									
3151	A A T T T A G A A A A A T T A T C T G G C T T T A C A A C A C T T A T T T G A T C A A C C T T A T A	515_al2.seq									
3150	A A T T T A G A A A A A T T A T C T G G C T T T A C A A C A C T T A T T T G A T C A A C C T T A T A	cjb111_al2.seq									
3150	A A T T T A G A A A A A T T A T C T G G C T T T A C A A C A C T T A T T T G A T C A A C C T T A T A	h36b_al2.seq									
		G G C T A A A A C T T C T T T G A T A T T A T G A A T A T A A A A A T T T T T C C T T T T T T A A Majority									
		3210		3220		3230		3240		3250	
3200	G G C T A A A A C T T C T T T G A T A T T A T G A A T A T A A A A A A T T T T T C C T T T T T T A A	2603_al2.seq									
3200	G G C T A A A A C T T C T T T G A T A T T A T G A A T A T A A A A A A T T T T T C C T T T T T T A A	18rs21_al2.seq									
3201	G G C T A A A A C T T C T T T G A T A T T A T G A A T A T A A A A A A T T T T T C C T T T T T T A A	515_al2.seq									
3200	G G C T A A A A C T T C T T T G A T A T T A T G A A T A T A A A A A A T T T T T C C T T T T T T A A	cjb111_al2.seq									
3200	G G C T A A A A C T T C T T T G A T A T T A T G A A T A T A A A A A A T T T T T C C T T T T T T A A	h36b_al2.seq									
		G T T T A T C T A A A T C T G T A A A T A A C T T A G C T T T A G G T A A G C C G C G A T G A G C T Majority									
		3260		3270		3280		3290		3300	
3250	G T T T A T C T A A A T C T G T A A A T A A C T T A G C T T T A G G T A A G C C G C G A T G A G C T	2603_al2.seq									
3250	G T T T A T C T A A A T C T G T A A A T A A C T T A G C T T T A G G T A A G C C G C G A T G A G C T	18rs21_al2.seq									
3251	G T T T A T C T A A A T C T G T A A A T A A C T T A G C T T T A G G T A A G C C G C G A T G A G C T	515_al2.seq									
3250	G T T T A T C T A A A T C T G T A A A T A A C T T A G C T T T A G G T A A G C C G C G A T G A G C T	cjb111_al2.seq									
3250	G T T T A T C T A A A T C T G T A A A T A A C T T A G C T T T A G G T A A G C C G C G A T G A G C T	h36b_al2.seq									
		G T G A T A A C A G T A T G T G A A C T T T T T C C A C C A A T T G G C A A G G A G G T T C C T T C Majority									
		3310		3320		3330		3340		3350	
3300	G T G A T A A C A G T A T G T G A A C T T T T T C C A C C A A T T G G C A A G G A G G T T C C T T C	2603_al2.seq									
3300	G T G A T A A C A G T A T G T G A A C T T T T T C C A C C A A T T G G C A A G G A G G T T C C T T C	18rs21_al2.seq									
3301	G T G A T A A C A G T A T G T G A A C T T T T T C C A C C A A T T G G C A A G G A G G T T C C T T C	515_al2.seq									
3300	G T G A T A A C A G T A T G T G A A C T T T T T C C A C C A A T T G G C A A G G A G G T T C C T T C	cjb111_al2.seq									
3300	G T G A T A A C A G T A T G T G A A C T T T T T C C A C C A A T T G G C A A G G A G G T T C C T T C	h36b_al2.seq									
		A A G G T G T C C T G C T C C T T T T T C A A G A A C A C T A C T G G T A G T C C C C G C A T A G A Majority									
		3360		3370		3380		3390		3400	
3350	A A G G T G T C C T G C T C C T T T T T C A A G A A C A C T A C T G G T A G T C C C C G C A T A G A	2603_al2.seq									
3350	A A G G T G T C C T G C T C C T T T T T C A A G A A C A C T A C T G G T A G T C C C C G C A T A G A	18rs21_al2.seq									
3351	A A G G T G T C C T G C T C C T T T T T C A A G A A C A C T A C T G G T A G T C C C C G C A T A G A	515_al2.seq									
3350	A A G G T G T C C T G C T C C T T T T T C A A G A A C A C T A C T G G T A G T C C C C G C A T A G A	cjb111_al2.seq									
3350	A A G G T G T C C T G C T C C T T T T T C A A G A A C A C T A C T G G T A G T C C C C G C A T A G A	h36b_al2.seq									
		T A G G T A A T T T T T G C T T G A T A G A C C G T A T A T C A A T A T A T C C A A T C A T T T C A Majority									
		3410		3420		3430		3440		3450	
3400	T A G G T A A T T T T T G C T T G A T A G A C C G T A T A T C A A T A T A T C C A A T C A T T T C A	2603_al2.seq									
3400	T A G G T A A T T T T T G C T T G A T A G A C C G T A T A T C A A T A T A T C C A A T C A T T T C A	18rs21_al2.seq									
3401	T A G G T A A T T T T T G C T T G A T A G A C C G T A T A T C A A T A T A T C C A A T C A T T T C A	515_al2.seq									
3400	T A G G T A A T T T T T G C T T G A T A G A C C G T A T A T C A A T A T A T C C A A T C A T T T C A	cjb111_al2.seq									
3400	T A G G T A A T T T T T G C T T G A T A G A C C G T A T A T C A A T A T A T C C A A T C A T T T C A	h36b_al2.seq									
		G C A A T C T C A A G C A T G T G G G C G T A T T C A G C A A T A C C T T T T T T T C T T T T T C Majority									
		3460		3470		3480		3490		3500	
3450	G C A A T C T C A A G C A T G T G G G C G T A T T C A G C A A T A C C T T T T T T T C T T T T T C	2603_al2.seq									
3450	G C A A T C T C A A G C A T G T G G G C G T A T T C A G C A A T A C C T T T T T T T C T T T T T C	18rs21_al2.seq									
3451	G C A A T C T C A A G C A T G T G G G C G T A T T C A G C A A T A C C T T T T T T T C T T T T T C	515_al2.seq									
3450	G C A A T C T C A A G C A T G T G G G C G T A T T C A G C A A T A C C T T T T T T T C T T T T T C	cjb111_al2.seq									
3450	G C A A T C T C A A G C A T G T G G G C G T A T T C A G C A A T A C C T T T T T T T C T T T T T C	h36b_al2.seq									
		A G T A T A G G G A T C T G A T A G G C G G C T T G G G T C C A G T G T T C T A T T A T A A G C T T Majority									
		3510		3520		3530		3540		3550	
3500	A G T A T A G G G A T C T G A T A G G C G G C T T G G G T C C A G T G T T C T A T T A T A A G C T T	2603_al2.seq									
3500	A G T A T A G G G A T C T G A T A G G C G G C T T G G G T C C A G T G T T C T A T T A T A A G C T T	18rs21_al2.seq									
3501	A G T A T A G G G A T C T G A T A G G C G G C T T G G G T C C A G T G T T C T A T T A T A A G C T T	515_al2.seq									
3500	A G T A T A G G G A T C T G A T A G G C G G C T T G G G T C C A G T G T T C T A T T A T A A G C T T	cjb111_al2.seq									
3500	A G T A T A G G G A T C T G A T A G G C G G C T T G G G T C C A G T G T T C T A T T A T A A G C T T	h36b_al2.seq									
		T T G C T A A C T C A A A T C G T C T A T T A A T C T C T T T A G T A T T T A A T T T T T G G G T T Majority									
		3560		3570		3580		3590		3600	
3550	T T G C T A A C T C A A A T C G T C T A T T A A T C T C T T T A G T A T T T A A T T T T T G G G T T	2603_al2.seq									
3550	T T G C T A A C T C A A A T C G T C T A T T A A T C T C T T T A G T A T T T A A T T T T T G G G T T	18rs21_al2.seq									
3551	T T G C T A A C T C A A A T C G T C T A T T A A T C T C T T T A G T A T T T A A T T T T T G G G T T	515_al2.seq									
3550	T T G C T A A C T C A A A T C G T C T A T T A A T C T C T T T A G T A T T T A A T T T T T G G G T T	cjb111_al2.seq									
3550	T T G C T A A C T C A A A T C G T C T A T T A A T C T C T T T A G T A T T T A A T T T T T G G G T T	h36b_al2.seq									

FIGURE 19C

		Majority									
		T G A T T A T C A A A G T T A G C T T A A T A T T A T A G T A C C A									
		3610 3620 3630 3640 3650									
3600	T G A T T A T C A A A G T T A G C T T A A T A T T A T A G T A C C A	2603_al2.seq									
3600	T G A T T A T C A A A G T T A G C T T A A T A T T A T A G T A C C A	18rs21_al2.seq									
3601	T G A T T A T C A A A G T T A G C T T A A T A T T A T A G T A C C A	515_al2.seq									
3600	T G A T T A T C A A A G T T A G C T T A A T A T T A T A G T A C C A	cjb111_al2.seq									
3600	T G A T T A T C A A A G T T A G C T T A A T A T T A T A G T A C C A	h36b_al2.seq									
		A T T T G A A A T A A A A G G A T A T G A G G T T A T C A A A A G A C C A A C T A A G A A C A A T A									
		3660 3670 3680 3690 3700									
3650	A T T T G A A A T A A A A G G A T A T G A G G T T A T C A A A A G A C C A A C T A A G A A C A A T A	2603_al2.seq									
3650	A T T T G A A A T A A A A G G A T A T G A G G T T A T C A A A A G A C C A A C T A A G A A C A A T A	18rs21_al2.seq									
3651	A T T T G A A A T A A A A G G A T A T G A G G T T A T C A A A A G A C C A A C T A A G A A C A A T A	515_al2.seq									
3650	A T T T G A A A T A A A A G G A T A T G A G G T T A T C A A A A G A C C A A C T A A G A A C A A T A	cjb111_al2.seq									
3650	A T T T G A A A T A A A A G G A T A T G A G G T T A T C A A A A G A C C A A C T A A G A A C A A T A	h36b_al2.seq									
		G T A T C A A G C C T A C A T T C A T C C A T C G A T T T A A A A C G A C C G A T T T C T T A A G G									
		3710 3720 3730 3740 3750									
3700	G T A T C A A G C C T A C A T T C A T C C A T C G A T T T A A A A C G A C C G A T T T C T T A A G G	2603_al2.seq									
3700	G T A T C A A G C C T A C A T T C A T C C A T C G A T T T A A A A C G A C C G A T T T C T T A A G G	18rs21_al2.seq									
3701	G T A T C A A G C C T A C A T T C A T C C A T C G A T T T A A A A C G A C C G A T T T C T T A A G G	515_al2.seq									
3700	G T A T C A A G C C T A C A T T C A T C C A T C G A T T T A A A A C G A C C G A T T T C T T A A G G	cjb111_al2.seq									
3700	G T A T C A A G C C T A C A T T C A T C C A T C G A T T T A A A A C G A C C G A T T T C T T A A G G	h36b_al2.seq									
		T T T T T C T G A A A T T T T C C T C C C A T T A T G A T T C A A T T C T T T T C T A A C A C T T									
		3760 3770 3780 3790 3800									
3750	T T T T T C T G A A A T T T T C C T C C C A T T A T G A T T C A A T T C T T T T C T A A C A C T T	2603_al2.seq									
3750	T T T T T C T G A A A T T T T C C T C C C A T T A T G A T T C A A T T C T T T T C T A A C A C T T	18rs21_al2.seq									
3751	T T T T T C T G A A A T T T T C C T C C C A T T A T G A T T C A A T T C T T T T C T A A C A C T T	515_al2.seq									
3750	T T T T T C T G A A A T T T T C C T C C C A T T A T G A T T C A A T T C T T T T C T A A C A C T T	cjb111_al2.seq									
3750	T T T T T C T G A A A T T T T C C T C C C A T T A T G A T T C A A T T C T T T T C T A A C A C T T	h36b_al2.seq									
		G C T A A A C G A T T T T T T T G A C G T T G A C G T T T T A T T A A C C A A A G T A A C C A A G C									
		3810 3820 3830 3840 3850									
3800	G C T A A A C G A T T T T T T T G A C G T T G A C G T T T T A T T A A C C A A A G T A A C C A A G C	2603_al2.seq									
3800	G C T A A A C G A T T T T T T T G A C G T T G A C G T T T T A T T A A C C A A A G T A A C C A A G C	18rs21_al2.seq									
3801	G C T A A A C G A T T T T T T T G A C G T T G A C G T T T T A T T A A C C A A A G T A A C C A A G C	515_al2.seq									
3800	G C T A A A C G A T T T T T T T G A C G T T G A C G T T T T A T T A A C C A A A G T A A C C A A G C	cjb111_al2.seq									
3800	G C T A A A C G A T T T T T T T G A C G T T G A C G T T T T A T T A A C C A A A G T A A C C A A G C	h36b_al2.seq									
		A A T A A T A A C T A A A G A T A T A T A G A A A A G A T A T C T A T A A A T C G T G T T T A A A T									
		3860 3870 3880 3890 3900									
3850	A A T A A T A A C T A A A G A T A T A T A G A A A A G A T A T C T A T A A A T C G T G T T T A A A T	2603_al2.seq									
3850	A A T A A T A A C T A A A G A T A T A T A G A A A A G A T A T C T A T A A A T C G T G T T T A A A T	18rs21_al2.seq									
3851	A A T A A T A A C T A A A G A T A T A T A G A A A A G A T A T C T A T A A A T C G T G T T T A A A T	515_al2.seq									
3850	A A T A A T A A C T A A A G A T A T A T A G A A A A G A T A T C T A T A A A T C G T G T T T A A A T	cjb111_al2.seq									
3850	A A T A A T A A C T A A A G A T A T A T A G A A A A G A T A T C T A T A A A T C G T G T T T A A A T	h36b_al2.seq									
		G A C C G T C T T T T A T T A A T T T T T C A T C A A T A G G A C C T T T A T A A G G G A T A C C A									
		3910 3920 3930 3940 3950									
3900	G A C C G T C T T T T A T T A A T T T T T C A T C A A T A G G A C C T T T A T A A G G G A T A C C A	2603_al2.seq									
3900	G A C C G T C T T T T A T T A A T T T T T C A T C A A T A G G A C C T T T A T A A G G G A T A C C A	18rs21_al2.seq									
3901	G A C C G T C T T T T A T T A A T T T T T C A T C A A T A G G A C C T T T A T A A G G G A T A C C A	515_al2.seq									
3900	G A C C G T C T T T T A T T A A T T T T T C A T C A A T A G G A C C T T T A T A A G G G A T A C C A	cjb111_al2.seq									
3900	G A C C G T C T T T T A T T A A T T T T T C A T C A A T A G G A C C T T T A T A A G G G A T A C C A	h36b_al2.seq									
		T G T C C C C T T A C T A A A A G T C T G T G T G T A T T G A C A T A A T C G G G G T G C A A G T									
		3960 3970 3980 3990 4000									
3950	T G T C C C C T T A C T A A A A G T C T G T G T G T A T T G A C A T A A T C G G G G T G C A A G T	2603_al2.seq									
3950	T G T C C C C T T A C T A A A A G T C T G T G T G T A T T G A C A T A A T C G G G G T G C A A G T	18rs21_al2.seq									
3951	T G T C C C C T T A C T A A A A G T C T G T G T G T A T T G A C A T A A T C G G G G T G C A A G T	515_al2.seq									
3950	T G T C C C C T T A C T A A A A G T C T G T G T G T A T T G A C A T A A T C G G G G T G C A A G T	cjb111_al2.seq									
3950	T G T C C C C T T A C T A A A A G T C T G T G T G T A T T G A C A T A A T C G G G G T G C A A G T	h36b_al2.seq									
		C A A T A A G G T T G C A T A A T C A T G T C C A G G A A C A A C C A A C A A T C T G A A A A G T									
		4010 4020 4030 4040 4050									
4000	C A A T A A G G T T G C A T A A T C A T G T C C A G G A A C A A C C A A C A A T C T G A A A A G T	2603_al2.seq									
4000	C A A T A A G G T T G C A T A A T C A T G T C C A G G A A C A A C C A A C A A T C T G A A A A G T	18rs21_al2.seq									
4001	C A A T A A G G T T G C A T A A T C A T G T C C A G G A A C A A C C A A C A A T C T G A A A A G T	515_al2.seq									
4000	C A A T A A G G T T G C A T A A T C A T G T C C A G G A A C A A C C A A C A A T C T G A A A A G T	cjb111_al2.seq									
4000	C A A T A A G G T T G C A T A A T C A T G T C C A G G A A C A A C C A A C A A T C T G A A A A G T	h36b_al2.seq									

FIGURE 19H

		TATCGGGCTGTTAAACGAGCTTTTATCTCTCCTTTTCTTAACTTTTAAAGGTT Majority									
		4060		4070		4080		4090		4100	
4050	TATCGGGCTGTTAAACGAGCTTTTATCTCTCCTTTTCTTAACTTTTAAAGGTT	2603_a12.seq									
4050	TATCGGGCTGTTAAACGAGCTTTTATCTCTCCTTTTCTTAACTTTTAAAGG										

66/487

AATCCGCTGCTTTGATTATTAGATTCTATCGTATAGTAAAAACGTGATACC Majority									
4510		4520		4530		4540			

67/487

C T T T T C T C T T T T G A I O R C T T T T G A T T T A A G T T T C A A A C T T A A C A T C A C C T G C C A A T Majority																																																	
4960										4970										4980										4990										5000									
4944	C	T	T	T	A	T	G	T	A	T	G	A	C	C	T	T	T	G	A	T	T	A	A	A	C	T	T	A	A	T	A	T	A	C	C	T	G	C	C	A	A	C	2603_al2.seq						
4944	C	T	T	T	A	T	G	T	A	T	A	C	C	T	T	T	G	A	T	T	A	A	A	C	T	T	A	A	T	A	T	A	C	C	T	G	C	C	A	A	C	18rs21_al2.seq							
4948	C	T	T	T	A	C	T	G	T	A	T	A	G	T	A	G	C	A	T	T	A	A	C	A	T	T	A	A	C	A	T	A	C	A	C	T	G	A	T	A	515_al2.seq								
4947	C	T	T	T	G	C	T	A	T	A	T	A	G	T	A	G	C	A	T	T	A	A	C	A	T	T	A	A	C	A	T	A	C	A	C	T	G	A	C	A	A	cjb111_al2.seq							
4944	C	G	T	T	T	G	C	T	A	T	A	G	C	T	T	T	G	A	A	A	T	T	A	A	C	A	T	T	A	A	C	A	T	A	C	C	T	G	C	T	A	A	h36b_al2.seq						
T T A G C A T A A C C T G C T G G A G C T T T G T T T T C T T C A A G G T T G T A A G T A C C T T C Majority																																																	
5010										5020										5030										5040										5050									
4994	T	T	A	G	C	A	A	A	T	C	T	G	C	T	G	G	A	G	C	T	T	T	C	T	T	C	A	A	G	G	T	T	G	T	A	A	G	T	A	C	C	G	T	C	2603_al2.seq				
4994	T	T	A	G	C	A	A	A	T	C	T	G	C	T	G	G	A	A	G	T	T	T	C	T	T	C	A	A	G	G	T	T	G	T	A	A	G	T	A	C	C	G	T	C	18rs21_al2.seq				
4998	T	T	A	G	C	A	T	A	A	C	C	A	G	T	T	G	G	A	G	C	T	T	T	C	T	C	A	A	T	A	G	T	A	T	T	A	C	C	T	T	C	515_al2.seq							
4997	T	T	T	G	C	A	T	A	A	C	C	T	G	T	G	G	A	G	C	T	T	T	C	T	C	A	A	T	A	G	T	A	T	T	A	C	C	T	T	T	cjb111_al2.seq								
4994	T	T	A	G	C	A	T	A	A	C	C	A	G	C	T	G	G	A	A	G	C	T	T	T	C	A	A	G	C	T	T	A	A	G	T	A	C	C	A	G	T	h36b_al2.seq							
T G C A A G A C C T G T A A C T T C A A A T T G A C C T T G A T C G T T T G A A G T C A A T T T A A Majority																																																	
5060										5070										5080										5090										5100									
5044	T	G	C	A	A	G	A	C	C	T	G	T	A	A	C	T	T	C	A	A	A	T	T	G	A	C	C	T	T	G	A	T	C	G	T	T	G	A	A	G	T	C	A	A	2603_al2.seq				
5044	T	G	C	A	A	G	A	C	C	T	G	T	A	A	C	T	T	C	A	A	A	T	T	G	A	C	C	T	T	G	A	T	C	G	T	T	G	A	A	G	T	C	A	A	18rs21_al2.seq				
5048	A	G	T	C	A	A	G	C	C	A	G	T	A	A	T	T	C	A	A	A	T	T	G	A	C	C	T	T	T	A	A	T	A	C	A	A	T	T	A	A	515_al2.seq								
5047	A	T	C	C	A	A	C	C	A	G	T	A	A	T	T	C	A	A	A	T	T	G	A	C	C	C	C	G	T	T	A	G	A	T	C	A	A	T	T	A	A	cjb111_al2.seq							
5044	T	G	C	A	A	G	A	C	C	T	G	T	A	A	C	T	T	C	A	A	A	T	T	G	A	C	C	T	T	G	A	A	G	T	C	A	T	T	T	A	A	h36b_al2.seq							
C A A C T T T A G C A G C - T T T T T A T C T G T T A C C C A C T C A T A A G C T G T A C G A G C T Majority																																																	
5110										5120										5130										5140										5150									
5094	T	G	G	C	T	T	A	G	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2603_al2.seq					
5094	T	G	G	C	T	T	A	G	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	18rs21_al2.seq					
5098	C	A	A	C	A	T	T	T	A	G	C	A	T	T	T	T	A	T	C	T	G	T	T	A	C	C	A	T	T	C	A	T	A	A	G	C	T	G	T	A	C	G	A	G	C	515_al2.seq			
5097	C	A	A	C	A	T	T	A	G	C	C	T	T	T	T	A	T	C	T	G	C	A	A	C	C	A	T	C	A	T	A	A	G	C	T	G	T	A	C	G	A	G	C	cjb111_al2.seq					
5094	C	A	A	C	G	T	T	A	G	C	A	T	T	T	T	A	T	A	G	T	T	A	C	C	A	C	T	C	T	A	A	G	C	T	G	T	A	C	G	A	G	C	h36b_al2.seq						
T C A A T G A A G G C T G C A T T G T A A G C T T C T T G T T T A G T T T T G A T A G T T G C T T T Majority																																																	
5160										5170										5180										5190										5200									
5138	T	C	A	A	T	G	A	A	G	G	C	T	G	C	A	T	T	G	T	T	A	A	T	T	C	T	T	T	A	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	2603_al2.seq				
5138	T	C	A	A	T	G	A	A	G	G	C	T	G	C	A	T	T	G	T	T	A	A	T	T	C	T	T	T	T	A	G	T	T	T	T	T	T	T	T	T	T	T	T	T	18rs21_al2.seq				
5148	T	T	A	A	C	A	A	A	G	G	C	A	T	T	T	G	T	A	A	G	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	515_al2.seq					
5147	T	T	A	A	C	A	A	A	G	G	C	A	T	T	T	G	T	A	A	G	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	cjb111_al2.seq					
5144	A	C	G	A	T	G	A	A	G	G	C	T	G	C	A	T	T	G	T	A	A	G	C	T	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	h36b_al2.seq					
T G C T G T T T T A C C T T T T T G A C C T T T T T G T T C T T C T G C A G A C A A C T T G T T A T Majority																																																	
5210										5220										5230										5240										5250									
5188	T	G	C	A	G	T	A	A	T	T	C	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	2603_al2.seq					
5188	T	G	C	A	G	T	A	A	T	T	C	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	18rs21_al2.seq					
5198	T	G	C	T	G	T	T	T	A	C	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	515_al2.seq					
5197	T	G	C	T	G	T	T	T	A	C	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	cjb111_al2.seq					
5194	A	G	C	T	G	T	T	T	A	C	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	h36b_al2.seq					
A A G C T G C G A T A G C T T C A T C T A A A G C T A T T T T A G T A G T A G C T A A A G C T G T T Majority																																																	
5260										5270										5280										5290										5300									
5238	A	A	G	C	T	G	C	G	A	T	A	G	C	T	T	C	A	T	C	T	A	A	A	G	C	T	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	2603_al2.seq					
5238	A	A	G	C	T	G	C	G	A	T	A	G	C	T	T	C	A	T	C	T	A	A	A	G	C	T	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	18rs21_al2.seq					
5248	A	A	G	C	T	T	T	A	A	G	C	A	G	C	A	T	C	A	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	515_al2.seq					
5247	A	A	G	C	T	T	T	A	A	G	C	A	G	C	A	T	C	A	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	cjb111_al2.seq					
5244	A	A	G	C	T	G	C	G	A	T	T	T	A	G	C	T	C	A	A	G	A	G	C	A	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	h36b_al2.seq					
T T T T G A G C T G C T T T T G C T T C A T C T G C T T T A A G T G C A A G G T A T T T A C C T G C Majority																																																	
5310										5320										5330										5340										5350									
5285	T	T	T	T	G	A	G	C	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2603_al2.seq					
5285	T	T	T	T	G	A	G	C	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	18rs21_al2.seq					
5298	T	T	T	T	T	G	C	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	515_al2.seq					
5297	T	T	T	T	T	T	G	C	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	cjb111_al2.seq					
5294	T	T	T	T	T	G	A	G	C	T	G	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	h36b_al2.seq					
T G A G T T T T T A A C A A C G A A T T G T G C A C C T G C C A A G C G T T C A G T T T C A G G T T Majority																																																	
5360										5370										5380										5390										5400									
5332	T	G	A	G	T	T	T	T	T	T	A	A	C	A	A	C	G	A	A	A	C	G	T	C	A	G	T	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2603_al2.seq					
5332	T	G	A	G	T	T	T	T	T	T	A	A	C	A	A	C	G	T	C	A	A	A	C	G	T	C	A	G	T	T	-	-	-	-	-	-	-	-	-	-	-	-	-	18rs21_al2.seq					
5348	T	-	-	-	-	T	T	T	T	A	A	C	A	A	A	G	A	A	G	A	A	G	A	A	G	A	A	G	A	A	G	A	A	G	A	A	G	A	A	G	A	A	515_al2.seq						
5347	T	-	-	-	-	T	T	T	T	T	A	A	C	A	A	A	G	A	A	G	A	A	G	A	A	G	A	A	G	A	A	G	A	A	G	A	A	G	A	A	h36b_al2.seq								
5344	T	G																																															

68/487

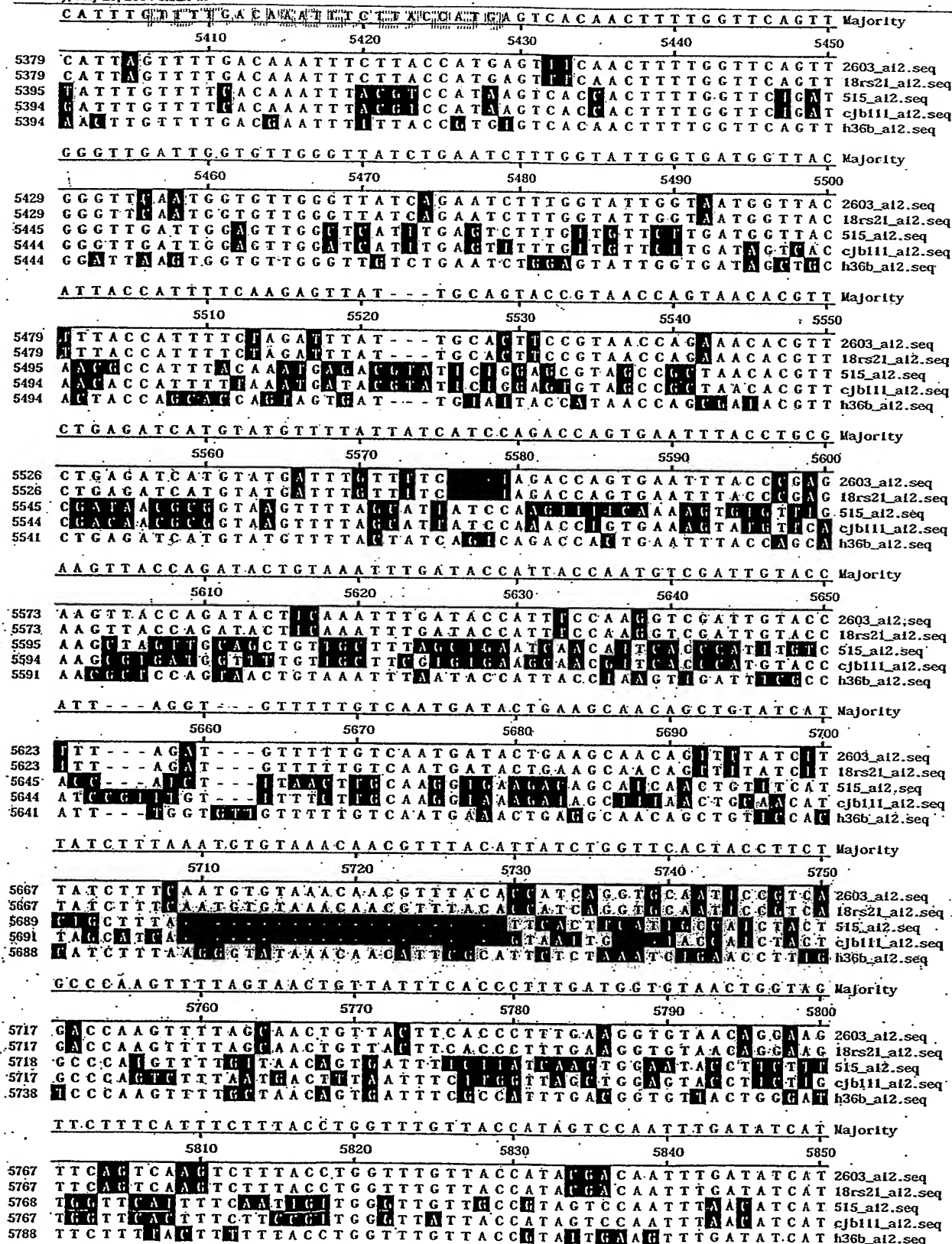


FIGURE 19L

69/487

T G C T T T C T G C A T T T T C A A C C A A A G C T T G C A C C G T T A A C A G T A G C T C T A T A A Majority									
5860		5870		5880		5890		5900	
5817	T G C A	T T C T C G A T T A	T C A A	T A A T	T G C T	T G A C C A	T T A A C A G T A G C A	A C T A T A A	2603_a12.seq
5817	T G C A	T T C T C G A T T A	T C A A	T A A T	T G C T	T G A C C A	T T A A C A G T A G C A	A C T A T A A	18rs21_a12.seq
5818	T G C T	T T C T C G A A C T	T T C A A C H A	T A G C A G A	A C C G T T	T A A G T A G C T C A G T A A			515_a12.seq
5817	T G C T	T T C T C G A A	T T T C A A C A G T A G	T G T G A G C C G T T	T A C C G T A G C T G A G T A A				cjb111_a12.seq
5838	T G C T	T T C T C G C T T T	T C A T C A A T A G C C G G	A C C G T T A A C T G T	T G C G C T A T A T				h36b_a12.seq
G T C A A T G T G A A T T C A A C A T C T T T T T G T T T T T G C T G C T T T T T C C A C T T T T G C Majority									
5910		5920		5930		5940		5950	
5867	G T C A A T G T	A A A T T C A A	T A T C A G G T	G T T T T A	G C T G C T	T T T T T T C C A A	T T T T G C		2603_a12.seq
5867	G T C A A T G T	A A A T T C A A	T A T C A G G T	G T T T T A	G C T G C T	T T T T T T C C A A	T T T T G C		18rs21_a12.seq
5868	G T H A I C I T	G A A T T C A A C A T C T T T T	T G T T T T T G C G G C T T T T	T C C A A T T T G C					515_a12.seq
5867	G T G A I C I T	G A A T T C A A C A T C T T T T	T G T T T T T G C G G C T T T T	T C C A A T T T G C					cjb111_a12.seq
5888	G T C A A T G T	G A A T T C A A C A G C T T	T G T T T T T G C T T	T C T T T T T G T A	G C T T A G C				h36b_a12.seq
C A A A C C T T T A G C T G T G A A T T T T A A T G T G A A A C C A C G G T C A T C A G T T C C A A Majority									
5960		5970		5980		5990		6000	
5917	C A A I C C A T	C A G C T G T G A A T T T T A A T G T G A A A C C A C G G	C A T C A A	T G C T A A					2603_a12.seq
5917	C A A T C C A T	C A G C T G T G A A T T T T A A T G T G A A A C C A C G G	C A T C A A	T G C T A A					18rs21_a12.seq
5918	A A G A C C T T	T T C A G T C A A A G A C A A G G C G G A A A C C	T T G T C A T C	T G T T A C A A					515_a12.seq
5917	A A G A C C T T	T T C A T T C A A G G A A A A G A A C C	T T G T C A T C	T G T T A C A A					cjb111_a12.seq
5938	C A A A C C T G	T A I C T G T G A A T T T T A A A G T G A A A C C A C G G T C A T C A G T T C C A A							h36b_a12.seq
G T T T G T A G T C T G T A T C C T T A A C A A A A G T T C C T G - - - - - C C G A T G C T T C A Majority									
6010		6020		6030		6040		6050	
5967	G T T T A T A G T C T G T A T C C T T A A	C A A A A G T T	T C T G A A G C T T	T A A					2603_a12.seq
5967	G T T T A T A G T C T G T A T C C T T A A	C A A A A G T T	T C T G A A G C T T	T A A					18rs21_a12.seq
5968	G T T T G T A A T	T G T A G C A T T A A G A G I I G A C C A I							515_a12.seq
5967	G T T T G T A G T	T A A A A C A G G A A A A I C I I A C C A T							cjb111_a12.seq
5988	C A T T T G T A A F C T G T A C	C H T T A C A A A A G A G C C T G							h36b_a12.seq
A G G T T A A C A G T G T T A C C C A T T G T C A A A C C A T T T G A C A T G C T A T C T G T C C A Majority									
6060		6070		6080		6090		6100	
6017	A G G T T A A C A G T	T G A A C C C A T T G T C A A A C C A T T T G A C A T	A T A T C T G T C C A						2603_a12.seq
6017	A G G T T A A C A G T	T G A A C C C A T T G T C A A A C C A T T T G A C A T	A T A T C T G T C C A						18rs21_a12.seq
6012									515_a12.seq
6011									cjb111_a12.seq
6032	A A G T T A A C A G T G T T A C C C A T T G T C A A A C C A T T T G A C A T G T A T C G A C C A								h36b_a12.seq
A A C C A A G T T T T T G T A T T T A G A A C C T T T T G T G A A T T T T T G T T T T A A C T T C G T Majority									
6110		6120		6130		6140		6150	
6067	A A C C A A G T T T T	T G T A T T T A G A A C C T T T T G T G A A T T T T T G T T T T A A C T T C A T							2603_a12.seq
6067	A A C C A A G T T T T	T G T A T T T A G A A C C T T T T G T G A A T T T T T G T T T T A A C T T C A T							18rs21_a12.seq
6059	G A I T A A	T T C T A T A G T C I G A A C C T T T A A G C A T T T T T T G T T C C A A C A T G G T							515_a12.seq
6058	A A C C A G T T T G T T A T A G T C I G A G C C T T T A A G A A T T T T T G T T C C A A C T A I G T								cjb111_a12.seq
6082	A G C A A G G T T T T T G T A T T T A G A A C C T T T T T G C A T T T T T G T T T T A A C T T C G T								h36b_a12.seq
A A G G T A T A A C T T T A C C G A T T T C A G C A G T A G C A G T T G C T T T G T C T T T T T G A Majority									
6160		6170		6180		6190		6200	
6117	A A G G A A A A A C T T T A C C G A T T T C A G C A G T A G C A G T T G C T T T G T C A C C T G C A								2603_a12.seq
6117	A A G G A A A A A C T T T A C C G A T T T C A G C A G T A G C A G T T G C T T T G T C A C C T G C A								18rs21_a12.seq
6109	A I I T I T A A C A T C A C C A A A G A A G C I G A G A C A G T T C T T T G T C T T T T T T G								515_a12.seq
6108	A I I C T I T T I I G T C A C C A A A T T I I G C I G A G A C A A A C C T T T G T C T T T T T T G								cjb111_a12.seq
6132	A A G G R A T T T C T T G A C C G A T T T C A G C A G T A G C I I T T G C T T T T T C T G C C T T G A								h36b_a12.seq
T A G T T A G C A T A A T C T G C G C C A G C T G T C A A A A G T C T A T T A A C A - C T G T C A A Majority									
6210		6220		6230		6240		6250	
6167	T A A T T A C C A T A A T	T T G C G C C A G C T G T C A A A A G T C T A T T A A C A							2603_a12.seq
6167	T A A T T A C C A T A A T	T T G C G C C A G C T G T C A A A A G T C T A T T A A C A							18rs21_a12.seq
6159	T T G T T G C C A T A A T C								515_a12.seq
6158	T T G T T A C T A T A A T C								cjb111_a12.seq
6182	T A T T T T G C T A T A T C T G C G C C A C G T G C A A A A G T T G A A T T A T C A - C T G T C A I								h36b_a12.seq
T - - T G T C A A G T T T G T T T T T A G C A A A G T T T T T A T C T A T T T G T G G T T T T T Majority									
6260		6270		6280		6290		6300	
6217	T G T G T C A A A T	C T T T G T T T A G C A A A G T T T T A T C A A T T T	T G G T T T T						2603_a12.seq
6217	T G T G T C A A A T	C T T T G T T T A G C A A A G T T T T A T C A A T T T	T G G T T T T						18rs21_a12.seq
6173									515_a12.seq
6172									cjb111_a12.seq
6231	T - - C A A C A A T T T A	T T T G C T T T A G C A A A G T T T T T T C A A T T T G T G G T G C T T							h36b_a12.seq

FIGURE 19M

70/487

CTTCACTGTTCTTTTGGATAAACAATGCGCATCTTCAACAACACCATCTTCC Majority									
6310		6320		6330		6340		6350	
6267	CTTCACTGTTCTTTTGGATAAACAATGCGCATCTTCAACAACACCATCTTCC	2603_a12.seq							
6267	CTTCACTGTTCTTTTGGATAAACAATGCGCATCTTCAACAACACCATCTTCC	18rs21_a12.seq							
6214	TTTCACTGTTCTTTTGGATAAACAATGCGCATCTTCAACAACACCATCTTCC	515_a12.seq							
6213	TTTCACTGTTCTTTTGGATAAACAATGCGCATCTTCAACAACACCATCTTCC	cjb111_a12.seq							
6279	CTTCTGTAATCTTTTGGATAAACAATGCGCATCTTCAACAACACCATCTTCC	h36b_a12.seq							
TTTACCAATGGAAGAGTGATTTTAACTGGAAGTCTTTTGAATCAGCCAG Majority									
6360		6370		6380		6390		6400	
6317	TTTACCAATGGAAGAGTGATTTTAACTGGAAGTCTTTTGAATCAGCCAG	2603_a12.seq							
6317	TTTACCAATGGAAGAGTGATTTTAACTGGAAGTCTTTTGAATCAGCCAG	18rs21_a12.seq							
6264	TTTACCAATGGAAGAGTGATTTTAACTGGAAGTCTTTTGAATCAGCCAG	515_a12.seq							
6263	TTTACCAATGGAAGAGTGATTTTAACTGGAAGTCTTTTGAATCAGCCAG	cjb111_a12.seq							
6329	TTTAAACAAGGAAGAGTGATTTTAACTGGAAGTCTTTTGAATCAGCCAG	h36b_a12.seq							
CATGGAACCATTTATTGTTGTAAGTAGATTTTCTTTCAATTCAACGATTT Majority									
6410		6420		6430		6440		6450	
6367	CATGGAACCATTTATTGTTGTAAGTAGATTTTCTTTCAATTCAACGATTT	2603_a12.seq							
6367	CATGGAACCATTTATTGTTGTAAGTAGATTTTCTTTCAATTCAACGATTT	18rs21_a12.seq							
6314	CATGGAACCATTTATTGTTGTAAGTAGATTTTCTTTCAATTCAACGATTT	515_a12.seq							
6313	CATGGAACCATTTATTGTTGTAAGTAGATTTTCTTTCAATTCAACGATTT	cjb111_a12.seq							
6379	AACTTTGTCATTTATTGTAAGTAGATTTTCTTTCAATTCAACGATTT	h36b_a12.seq							
GGTACTTTCCTTTTAAATTTTGTAGTGTAAAAACAA - - - - - TACCTTTT Majority									
6460		6470		6480		6490		6500	
6417	AACTTTGTCATTTATTGTAAGTAGATTTTCTTTCAATTCAACGATTT	2603_a12.seq							
6417	AACTTTGTCATTTATTGTAAGTAGATTTTCTTTCAATTCAACGATTT	18rs21_a12.seq							
6364	GGTAAAGTTTCTTTTAAATTTTGTAGTGTAAAAACAA - - - - - TACCTTTT	515_a12.seq							
6363	GGTAAAGTTTCTTTTAAATTTTGTAGTGTAAAAACAA - - - - - TACCTTTT	cjb111_a12.seq							
6429	GGTAAAGTTTCTTTTAAATTTTGTAGTGTAAAAACAA - - - - - TACCTTTT	h36b_a12.seq							
TCTGTTGTCATTC - - - - - TTTAAG Majority									
6510		6520		6530		6540		6550	
6467	TCTGTTGTCATTC - - - - - TTTAAG	2603_a12.seq							
6467	TCTGTTGTCATTC - - - - - TTTAAG	18rs21_a12.seq							
6408	TCTGTTGTCATTC - - - - - TTTAAG	515_a12.seq							
6407	TCTGTTGTCATTC - - - - - TTTAAG	cjb111_a12.seq							
6473	TCTGTTGTCATTC - - - - - TTTAAG	h36b_a12.seq							
AACACCATCTTTCATTAATTTCTGTTGGTTTTAAATTTGTTA - CTTCCTTAC Majority									
6560		6570		6580		6590		6600	
6517	AACACCATCTTTCATTAATTTCTGTTGGTTTTAAATTTGTTA - CTTCCTTAC	2603_a12.seq							
6517	AACACCATCTTTCATTAATTTCTGTTGGTTTTAAATTTGTTA - CTTCCTTAC	18rs21_a12.seq							
6428	AACACCATCTTTCATTAATTTCTGTTGGTTTTAAATTTGTTA - CTTCCTTAC	515_a12.seq							
6427	AACACCATCTTTCATTAATTTCTGTTGGTTTTAAATTTGTTA - CTTCCTTAC	cjb111_a12.seq							
6493	AACACCATCTTTCATTAATTTCTGTTGGTTTTAAATTTGTTA - CTTCCTTAC	h36b_a12.seq							
CACTTTCGGTAATGTAATTTTG - - - - - TTTTATCTTCATTTCATAAC - - - AGC Majority									
6610		6620		6630		6640		6650	
6566	CACTTTCGGTAATGTAATTTTG - - - - - TTTTATCTTCATTTCATAAC - - - AGC	2603_a12.seq							
6566	CACTTTCGGTAATGTAATTTTG - - - - - TTTTATCTTCATTTCATAAC - - - AGC	18rs21_a12.seq							
6478	CACTTTCGGTAATGTAATTTTG - - - - - TTTTATCTTCATTTCATAAC - - - AGC	515_a12.seq							
6477	CACTTTCGGTAATGTAATTTTG - - - - - TTTTATCTTCATTTCATAAC - - - AGC	cjb111_a12.seq							
6528	CACTTTCGGTAATGTAATTTTG - - - - - TTTTATCTTCATTTCATAAC - - - AGC	h36b_a12.seq							
AAAGAAAGCACCCTTCGATTTCTTTAGAT - - - CCTTCGCCAAAGTAACCTC Majority									
6660		6670		6680		6690		6700	
6616	AAAGAAAGCACCCTTCGATTTCTTTAGAT - - - CCTTCGCCAAAGTAACCTC	2603_a12.seq							
6616	AAAGAAAGCACCCTTCGATTTCTTTAGAT - - - CCTTCGCCAAAGTAACCTC	18rs21_a12.seq							
6521	AAAGAAAGCACCCTTCGATTTCTTTAGAT - - - CCTTCGCCAAAGTAACCTC	515_a12.seq							
6520	AAAGAAAGCACCCTTCGATTTCTTTAGAT - - - CCTTCGCCAAAGTAACCTC	cjb111_a12.seq							
6571	AAAGAAAGCACCCTTCGATTTCTTTAGAT - - - CCTTCGCCAAAGTAACCTC	h36b_a12.seq							
TAAGGTCAGTAATTTGTTTACCTATGTAGTCTTTTCCATTTCATACCTTTT Majority									
6710		6720		6730		6740		6750	
6663	TAAGGTCAGTAATTTGTTTACCTATGTAGTCTTTTCCATTTCATACCTTTT	2603_a12.seq							
6663	TAAGGTCAGTAATTTGTTTACCTATGTAGTCTTTTCCATTTCATACCTTTT	18rs21_a12.seq							
6571	TAAGGTCAGTAATTTGTTTACCTATGTAGTCTTTTCCATTTCATACCTTTT	515_a12.seq							
6570	TAAGGTCAGTAATTTGTTTACCTATGTAGTCTTTTCCATTTCATACCTTTT	cjb111_a12.seq							
6618	TAAGGTCAGTAATTTGTTTACCTATGTAGTCTTTTCCATTTCATACCTTTT	h36b_a12.seq							

FIGURE 19N

71/487

GTTCCTGGGAAGTTACTTTTGTTCAGAAATAGCTTGTGTTTGGACAATCTT Majority									
6760		6770		6780		6790		6800	
6713	GTTCCTGGGAAGTTACTTTTGT	1AAGATT	1GATT	1GGTTTG	1AAAATCTT	2603_a12.seq			
6713	GTTCCTGGGAAGTTACTTTTGT	1AAGATT	1GATT	1GGTTTG	1AAAATCTT	18rs21_a12.seq			
6621	GTACCAGCAG	1TAAAGTT	1TCAAA	1TGAAGTT	1GCAATCTT	515_a12.seq			
6620	GTACCTTCAG	1TAAAGTT	1TCAAA	1TGAAGTT	1GCAATCTT	cjb111_a12.seq			
6668	ATTCTCTGGGAATTTTGGTTT	1AAGTT	1TGAAGTT	1GCAATCTT	1TTC	h36b_a12.seq			
GTGCAAGGTCACTGTATTAGT---TGT---TGCTTCATCCGCAAAACG Majority									
6810		6820		6830		6840		6850	
6763	GTGCAAAGTCACTGTATTAGT	---TGT	---TGCTTCATCCGCAAAACG	2603_a12.seq					
6763	GTGCAAAGTCACTGTATTAGT	---TGT	---TGCTTCATCCGCAAAACG	18rs21_a12.seq					
6668	GTGCAAGGTCACTGTATTAGT	---TGT	---TGCTTCATCCGCAAAACG	515_a12.seq					
6667	GTGCAAGGTCACTGTATTAGT	---TGT	---TGCTTCATCCGCAAAACG	cjb111_a12.seq					
6718	GTGTAAGGTCACTGTATTAGT	---TGT	---TGCTTCATCCGCAAAACG	h36b_a12.seq					
CTGGTGCAACTGAGAACAAATGACGTTAAAGTCAGTAACAATGCCGAGAAAC Majority									
6860		6870		6880		6890		6900	
6804	CTGGTGCAACTGAGAACAAAT	GACGTTAAAGTCAGTAACAAT	GCCGAGAAAC	2603_a12.seq					
6804	CTGGTGCAACTGAGAACAAAT	GACGTTAAAGTCAGTAACAAT	GCCGAGAAAC	18rs21_a12.seq					
6712	CTGGTGCAACTGAGAACAAAT	GACGTTAAAGTCAGTAACAAT	GCCGAGAAAC	515_a12.seq					
6708	CTGGTGCAACTGAGAACAAAT	GACGTTAAAGTCAGTAACAAT	GCCGAGAAAC	cjb111_a12.seq					
6765	CTGGTGCAACTGAGAACAAAT	GACGTTAAAGTCAGTAACAAT	GCCGAGAAAC	h36b_a12.seq					
ATTGCAAAATATTTGTTGATTCTTTTCATTCTATCTCCTTCTTATTTTA Majority									
6910		6920		6930		6940		6950	
6854	ATTGCAAAATATTTGTTGATT	CTTTTCATTCTATCTCCTTCTTATTTTA	2603_a12.seq						
6854	ATTGCAAAATATTTGTTGATT	CTTTTCATTCTATCTCCTTCTTATTTTA	18rs21_a12.seq						
6762	ATTGCAAAATATTTGTTGATT	CTTTTCATTCTATCTCCTTCTTATTTTA	515_a12.seq						
6758	ATTGCAAAATATTTGTTGATT	CTTTTCATTCTATCTCCTTCTTATTTTA	cjb111_a12.seq						
6815	ATTGCAAAATATTTGTTGATT	CTTTTCATTCTATCTCCTTCTTATTTTA	h36b_a12.seq						
GTTAATCAACATGATTAATAATATGCGGATTTTAATA-C-ACCGCAGCAC Majority									
6960		6970		6980		6990		7000	
6904	GTTAATCAACATGATTAATAAT	ATGCGGATTTTAATA-C-ACCGCAGCAC	2603_a12.seq						
6904	GTTAATCAACATGATTAATAAT	ATGCGGATTTTAATA-C-ACCGCAGCAC	18rs21_a12.seq						
6812	GTTAATCAACATGATTAATAAT	ATGCGGATTTTAATA-C-ACCGCAGCAC	515_a12.seq						
6808	GTTAATCAACATGATTAATAAT	ATGCGGATTTTAATA-C-ACCGCAGCAC	cjb111_a12.seq						
6865	GTTAATCAACATGATTAATAAT	ATGCGGATTTTAATA-C-ACCGCAGCAC	h36b_a12.seq						
CACTCCCTTCAAGTCATGGAATTTTAGTTAATTAATTAAGAATACTAAAG Majority									
7010		7020		7030		7040		7050	
6952	CACTCCCTTCAAGTCATGGAAT	TTTAGTTAATTAATTAAGAATACTAAAG	2603_a12.seq						
6952	CACTCCCTTCAAGTCATGGAAT	TTTAGTTAATTAATTAAGAATACTAAAG	18rs21_a12.seq						
6860	CACTCCCTTCAAGTCATGGAAT	TTTAGTTAATTAATTAAGAATACTAAAG	515_a12.seq						
6857	CACTCCCTTCAAGTCATGGAAT	TTTAGTTAATTAATTAAGAATACTAAAG	cjb111_a12.seq						
6914	CACTCCCTTCAAGTCATGGAAT	TTTAGTTAATTAATTAAGAATACTAAAG	h36b_a12.seq						
CGCATAAATTTTAAATCTTTTTTGTATGGACATATCACTAGATTTCCTTATAC Majority									
7060		7070		7080		7090		7100	
7002	CGCATAAATTTTAAATCTTTT	TGTATGGACATATCACTAGATTTCCTTATAC	2603_a12.seq						
7002	CGCATAAATTTTAAATCTTTT	TGTATGGACATATCACTAGATTTCCTTATAC	18rs21_a12.seq						
6910	CGCATAAATTTTAAATCTTTT	TGTATGGACATATCACTAGATTTCCTTATAC	515_a12.seq						
6899	CGCATAAATTTTAAATCTTTT	TGTATGGACATATCACTAGATTTCCTTATAC	cjb111_a12.seq						
6952	CGCATAAATTTTAAATCTTTT	TGTATGGACATATCACTAGATTTCCTTATAC	h36b_a12.seq						
CTTTTCCAAATATAAAATTCACCTGCAATAGACATCATAGCTCCACCTAT Majority									
7110		7120		7130		7140		7150	
7052	CTTTTCCAAATATAAAATTCAC	CTGCAATAGACATCATAGCTCCACCTAT	2603_a12.seq						
7052	CTTTTCCAAATATAAAATTCAC	CTGCAATAGACATCATAGCTCCACCTAT	18rs21_a12.seq						
6960	CTTTTCCAAATATAAAATTCAC	CTGCAATAGACATCATAGCTCCACCTAT	515_a12.seq						
6949	CTTTTCCAAATATAAAATTCAC	CTGCAATAGACATCATAGCTCCACCTAT	cjb111_a12.seq						
7002	CTTTTCCAAATATAAAATTCAC	CTGCAATAGACATCATAGCTCCACCTAT	h36b_a12.seq						
TAAAAATGAAAGATAGAATTCTTTTCCACCTGTCATAGGAATAATTCCTT Majority									
7160		7170		7180		7190		7200	
7102	TAAAAATGAAAGATAGAATTCT	TTTCCACCTGTCATAGGAATAATTCCTT	2603_a12.seq						
7102	TAAAAATGAAAGATAGAATTCT	TTTCCACCTGTCATAGGAATAATTCCTT	18rs21_a12.seq						
7010	TAAAAATGAAAGATAGAATTCT	TTTCCACCTGTCATAGGAATAATTCCTT	515_a12.seq						
6999	TAAAAATGAAAGATAGAATTCT	TTTCCACCTGTCATAGGAATAATTCCTT	cjb111_a12.seq						
7052	TAAAAATGAAAGATAGAATTCT	TTTCCACCTGTCATAGGAATAATTCCTT	h36b_a12.seq						

FIGURE 190

72/487

	TTGGTGGGAATATGCGTGTGTTGGTAATTAATGCTTGTACCTTCCTCATGA Majority									
	7210		7220		7230		7240		7250	
7152	TTGGTGGGAATATGCGTGTGTTGGTAATTAATGCTTGTACCTTCCTCATGA 2603_a12.seq									
7152	TTGGTGGGAATATGCGTGTGTTGGTAATTAATGCTTGTACCTTCCTCATGA 18rs21_a12.seq									
7060	TTGGTGGGAATATGCGTGTGTTGGTAATTAATGCTTGTACCTTCCTCATGA 515_a12.seq									
7049	TTGGTGGGAATATGCGTGTGTTGGTAATTAATGCTTGTACCTTCCTCATGA cjb111_a12.seq									
7102	TTGGTGGGAATATGCGTGTGTTGGTAATTAATGCTTGTACCTTCCTCATGA h36b_a12.seq									
	TATTCAGAAATCTGTTTATTAACAGCTATTATATTTTTTATCGATCCTTT Majority									
	7260		7270		7280		7290		7300	
7202	TATTCAGAAATCTGTTTATTAACAGCTATTATATTTTTTATCGATCCTTT 2603_a12.seq									
7202	TATTCAGAAATCTGTTTATTAACAGCTATTATATTTTTTATCGATCCTTT 18rs21_a12.seq									
7110	TATTCAGAAATCTGTTTATTAACAGCTATTATATTTTTTATCGATCCTTT 515_a12.seq									
7099	TATTCAGAAATCTGTTTATTAACAGCTATTATATTTTTTATCGATCCTTT cjb111_a12.seq									
7152	TATTCAGAAATCTGTTTATTAACAGCTATTATATTTTTTATCGATCCTTT h36b_a12.seq									
	AACCACCTTCAAAAGTTAAATTTGGTTTATTAGTAATTTTTTGATAATCCT Majority									
	7310		7320		7330		7340		7350	
7252	AACCACCTTCAAAAGTTAAATTTGGTTTATTAGTAATTTTTTGATAATCCT 2603_a12.seq									
7252	AACCACCTTCAAAAGTTAAATTTGGTTTATTAGTAATTTTTTGATAATCCT 18rs21_a12.seq									
7160	AACCACCTTCAAAAGTTAAATTTGGTTTATTAGTAATTTTTTGATAATCCT 515_a12.seq									
7149	AACCACCTTCAAAAGTTAAATTTGGTTTATTAGTAATTTTTTGATAATCCT cjb111_a12.seq									
7202	AACCACCTTCAAAAGTTAAATTTGGTTTATTAGTAATTTTTTGATAATCCT h36b_a12.seq									
	CCGGCGAAACTGCTTCTATTAACCTGATATTTGCCATCTTTTCAAAATCTTTG Majority									
	7360		7370		7380		7390		7400	
7302	CCGGCGAAACTGCTTCTATTAACCTGATATTTGCCATCTTTTCAAAATCTTTG 2603_a12.seq									
7302	CCGGCGAAACTGCTTCTATTAACCTGATATTTGCCATCTTTTCAAAATCTTTG 18rs21_a12.seq									
7210	CCGGCGAAACTGCTTCTATTAACCTGATATTTGCCATCTTTTCAAAATCTTTG 515_a12.seq									
7199	CCGGCGAAACTGCTTCTATTAACCTGATATTTGCCATCTTTTCAAAATCTTTG cjb111_a12.seq									
7252	CCGGCGAAACTGCTTCTATTAACCTGATATTTGCCATCTTTTCAAAATCTTTG h36b_a12.seq									
	TAAGAAATTTTGGCGTTTCTCCCGTCACTACTTTTGAATTATTATTTTTT Majority									
	7410		7420		7430		7440		7450	
7352	TAAGAAATTTTGGCGTTTCTCCCGTCACTACTTTTGAATTATTATTTTTT 2603_a12.seq									
7352	TAAGAAATTTTGGCGTTTCTCCCGTCACTACTTTTGAATTATTATTTTTT 18rs21_a12.seq									
7260	TAAGAAATTTTGGCGTTTCTCCCGTCACTACTTTTGAATTATTATTTTTT 515_a12.seq									
7249	TAAGAAATTTTGGCGTTTCTCCCGTCACTACTTTTGAATTATTATTTTTT cjb111_a12.seq									
7302	TAAGAAATTTTGGCGTTTCTCCCGTCACTACTTTTGAATTATTATTTTTT h36b_a12.seq									
	TATTGGTAAATAAAGTTTATAATCTTCATTAAATTCCTTGAAGTTCAAACG Majority									
	7460		7470		7480		7490		7500	
7402	TATTGGTAAATAAAGTTTATAATCTTCATTAAATTCCTTGAAGTTCAAACG 2603_a12.seq									
7402	TATTGGTAAATAAAGTTTATAATCTTCATTAAATTCCTTGAAGTTCAAACG 18rs21_a12.seq									
7310	TATTGGTAAATAAAGTTTATAATCTTCATTAAATTCCTTGAAGTTCAAACG 515_a12.seq									
7299	TATTGGTAAATAAAGTTTATAATCTTCATTAAATTCCTTGAAGTTCAAACG cjb111_a12.seq									
7352	TATTGGTAAATAAAGTTTATAATCTTCATTAAATTCCTTGAAGTTCAAACG h36b_a12.seq									
	TAGCTCCTTTGAGAAGCAACTTATTATTATCTTTATCAACTTTTATAAAT Majority									
	7510		7520		7530		7540		7550	
7452	TAGCTCCTTTGAGAAGCAACTTATTATTATCTTTATCAACTTTTATAAAT 2603_a12.seq									
7452	TAGCTCCTTTGAGAAGCAACTTATTATTATCTTTATCAACTTTTATAAAT 18rs21_a12.seq									
7360	TAGCTCCTTTGAGAAGCAACTTATTATTATCTTTATCAACTTTTATAAAT 515_a12.seq									
7349	TAGCTCCTTTGAGAAGCAACTTATTATTATCTTTATCAACTTTTATAAAT cjb111_a12.seq									
7402	TAGCTCCTTTGAGAAGCAACTTATTATTATCTTTATCAACTTTTATAAAT h36b_a12.seq									
	TCAATTTTCACTTAACCTTCTTCTCGTTTTTAAATCGTTATTGTAGGATATTG Majority									
	7560		7570		7580		7590		7600	
7502	TCAATTTTCACTTAACCTTCTTCTCGTTTTTAAATCGTTATTGTAGGATATTG 2603_a12.seq									
7502	TCAATTTTCACTTAACCTTCTTCTCGTTTTTAAATCGTTATTGTAGGATATTG 18rs21_a12.seq									
7410	TCAATTTTCACTTAACCTTCTTCTCGTTTTTAAATCGTTATTGTAGGATATTG 515_a12.seq									
7399	TCAATTTTCACTTAACCTTCTTCTCGTTTTTAAATCGTTATTGTAGGATATTG cjb111_a12.seq									
7452	TCAATTTTCACTTAACCTTCTTCTCGTTTTTAAATCGTTATTGTAGGATATTG h36b_a12.seq									
	TCTCACATCACGAATTTTAGGGATTGCAAAATCTCTAAAGTGTATTAGGAT Majority									
	7610		7620		7630		7640		7650	
7552	TCTCACATCACGAATTTTAGGGATTGCAAAATCTCTAAAGTGTATTAGGAT 2603_a12.seq									
7552	TCTCACATCACGAATTTTAGGGATTGCAAAATCTCTAAAGTGTATTAGGAT 18rs21_a12.seq									
7460	TCTCACATCACGAATTTTAGGGATTGCAAAATCTCTAAAGTGTATTAGGAT 515_a12.seq									
7449	TCTCACATCACGAATTTTAGGGATTGCAAAATCTCTAAAGTGTATTAGGAT cjb111_a12.seq									
7502	TCTCACATCACGAATTTTAGGGATTGCAAAATCTCTAAAGTGTATTAGGAT h36b_a12.seq									

FIGURE 19P

73/487

	CCTCTGACTTAGGATTCAATGTTGTTCTACCATTAAGTGTTCATAGAATTTG	Majority
	7660 7670 7680 7690 7700	
7602	CCTCTGACTTAGGATTCAATGTTGTTCTACCATTAAGTGTTCATAGAATTTG	2603_a12.seq
7602	CCTCTGACTTAGGATTCAATGTTGTTCTACCATTAAGTGTTCATAGAATTTG	18rs21_a12.seq
7510	CCTCTGACTTAGGATTCAATGTTGTTCTACCATTAAGTGTTCATAGAATTTG	515_a12.seq
7499	CCTCTGACTTAGGATTCAATGTTGTTCTACCATTAAGTGTTCATAGAATTTG	cjb111_a12.seq
7552	CCTCTGACTTAGGATTCAATGTTGTTCTACCATTAAGTGTTCATAGAATTTG	h36b_a12.seq
	TTACTTTATAAAACTGTCATCTAGTTTTCACATCATATGTGAGTGTACTTT	Majority
	7710 7720 7730 7740 7750	
7652	TTACTTTATAAAACTGTCATCTAGTTTTCACATCATATGTGAGTGTACTTT	2603_a12.seq
7652	TTACTTTATAAAACTGTCATCTAGTTTTCACATCATATGTGAGTGTACTTT	18rs21_a12.seq
7560	TTACTTTATAAAACTGTCATCTAGTTTTCACATCATATGTGAGTGTACTTT	515_a12.seq
7549	TTACTTTATAAAACTGTCATCTAGTTTTCACATCATATGTGAGTGTACTTT	cjb111_a12.seq
7602	TTACTTTATAAAACTGTCATCTAGTTTTCACATCATATGTGAGTGTACTTT	h36b_a12.seq
	TTGACCTTCTCCTAAATTCAAAACCTCTAAACATAGAGTTTATTTCCGATGT	Majority
	7760 7770 7780 7790 7800	
7702	TTGACCTTCTCCTAAATTCAAAACCTCTAAACATAGAGTTTATTTCCGATGT	2603_a12.seq
7702	TTGACCTTCTCCTAAATTCAAAACCTCTAAACATAGAGTTTATTTCCGATGT	18rs21_a12.seq
7610	TTGACCTTCTCCTAAATTCAAAACCTCTAAACATAGAGTTTATTTCCGATGT	515_a12.seq
7599	TTGACCTTCTCCTAAATTCAAAACCTCTAAACATAGAGTTTATTTCCGATGT	cjb111_a12.seq
7652	TTGACCTTCTCCTAAATTCAAAACCTCTAAACATAGAGTTTATTTCCGATGT	h36b_a12.seq
	ATTCTAATTTAAACCCCTTAAGTATTCACCATCATTATTAGGCCCAACCA	Majority
	7810 7820 7830 7840 7850	
7752	ATTCTAATTTAAACCCCTTAAGTATTCACCATCATTATTAGGCCCAACCA	2603_a12.seq
7752	ATTCTAATTTAAACCCCTTAAGTATTCACCATCATTATTAGGCCCAACCA	18rs21_a12.seq
7660	ATTCTAATTTAAACCCCTTAAGTATTCACCATCATTATTAGGCCCAACCA	515_a12.seq
7649	ATTCTAATTTAAACCCCTTAAGTATTCACCATCATTATTAGGCCCAACCA	cjb111_a12.seq
7702	ATTCTAATTTAAACCCCTTAAGTATTCACCATCATTATTAGGCCCAACCA	h36b_a12.seq
	GTTGCAATACCATCCTTCATTACACTTCCATCATTTCCTGTAAAGTATA	Majority
	7860 7870 7880 7890 7900	
7802	GTTGCAATACCATCCTTCATTACACTTCCATCATTTCCTGTAAAGTATA	2603_a12.seq
7802	GTTGCAATACCATCCTTCATTACACTTCCATCATTTCCTGTAAAGTATA	18rs21_a12.seq
7710	GTTGCAATACCATCCTTCATTACACTTCCATCATTTCCTGTAAAGTATA	515_a12.seq
7699	GTTGCAATACCATCCTTCATTACACTTCCATCATTTCCTGTAAAGTATA	cjb111_a12.seq
7752	GTTGCAATACCATCCTTCATTACACTTCCATCATTTCCTGTAAAGTATA	h36b_a12.seq
	ATCACTTGGCTGTAAATGTTTGTCCATTACCAAGCTGTAAATTGATTTTAT	Majority
	7910 7920 7930 7940 7950	
7852	ATCACTTGGCTGTAAATGTTTGTCCATTACCAAGCTGTAAATTGATTTTAT	2603_a12.seq
7852	ATCACTTGGCTGTAAATGTTTGTCCATTACCAAGCTGTAAATTGATTTTAT	18rs21_a12.seq
7760	ATCACTTGGCTGTAAATGTTTGTCCATTACCAAGCTGTAAATTGATTTTAT	515_a12.seq
7749	ATCACTTGGCTGTAAATGTTTGTCCATTACCAAGCTGTAAATTGATTTTAT	cjb111_a12.seq
7802	ATCACTTGGCTGTAAATGTTTGTCCATTACCAAGCTGTAAATTGATTTTAT	h36b_a12.seq
	CACCCATAGGATCTTCGATAGTTCCATTAAACAATTGAGTTTCTTTTGT	Majority
	7960 7970 7980 7990 8000	
7902	CACCCATAGGATCTTCGATAGTTCCATTAAACAATTGAGTTTCTTTTGT	2603_a12.seq
7902	CACCCATAGGATCTTCGATAGTTCCATTAAACAATTGAGTTTCTTTTGT	18rs21_a12.seq
7810	CACCCATAGGATCTTCGATAGTTCCATTAAACAATTGAGTTTCTTTTGT	515_a12.seq
7799	CACCCATAGGATCTTCGATAGTTCCATTAAACAATTGAGTTTCTTTTGT	cjb111_a12.seq
7852	CACCCATAGGATCTTCGATAGTTCCATTAAACAATTGAGTTTCTTTTGT	h36b_a12.seq
	AAAATCGTTTCAAATTTGTTGCTGAATTTTAGATAAAAATTTCAATTGTTAGA	Majority
	8010 8020 8030 8040 8050	
7952	AAAATCGTTTCAAATTTGTTGCTGAATTTTAGATAAAAATTTCAATTGTTAGA	2603_a12.seq
7952	AAAATCGTTTCAAATTTGTTGCTGAATTTTAGATAAAAATTTCAATTGTTAGA	18rs21_a12.seq
7860	AAAATCGTTTCAAATTTGTTGCTGAATTTTAGATAAAAATTTCAATTGTTAGA	515_a12.seq
7849	AAAATCGTTTCAAATTTGTTGCTGAATTTTAGATAAAAATTTCAATTGTTAGA	cjb111_a12.seq
7902	AAAATCGTTTCAAATTTGTTGCTGAATTTTAGATAAAAATTTCAATTGTTAGA	h36b_a12.seq
	TGTATCGGCTGAAGTTACGATAGGGGTGTAGTACTCAGCTTTTGAAGAGA	Majority
	8060 8070 8080 8090 8100	
8002	TGTATCGGCTGAAGTTACGATAGGGGTGTAGTACTCAGCTTTTGAAGAGA	2603_a12.seq
8002	TGTATCGGCTGAAGTTACGATAGGGGTGTAGTACTCAGCTTTTGAAGAGA	18rs21_a12.seq
7910	TGTATCGGCTGAAGTTACGATAGGGGTGTAGTACTCAGCTTTTGAAGAGA	515_a12.seq
7899	TGTATCGGCTGAAGTTACGATAGGGGTGTAGTACTCAGCTTTTGAAGAGA	cjb111_a12.seq
7952	TGTATCGGCTGAAGTTACGATAGGGGTGTAGTACTCAGCTTTTGAAGAGA	h36b_a12.seq

FIGURE 19Q

74/487

	8110	8120	8130	8140	8150	
	ACGACCTCATTAGTTCTGTGATTTCTCCATCTGAAAGTTTAAAAGCTTCC Majority					
8052	ACGACCTCATTAGTTCTGTGATTTCTCCATCTGAAAGTTTAAAAGCTTCC					2603_al2.seq
8052	ACGACCTCATTAGTTCTGTGATTTCTCCATCTGAAAGTTTAAAAGCTTCC					18rs21_al2.seq
7960	ACGACCTCATTAGTTCTGTGATTTCTCCATCTGAAAGTTTAAAAGCTTCC					515_al2.seq
7949	ACGACCTCATTAGTTCTGTGATTTCTCCATCTGAAAGTTTAAAAGCTTCC					cjb111_al2.seq
8002	ACGACCTCATTAGTTCTGTGATTTCTCCATCTGAAAGTTTAAAAGCTTCC					h36b_al2.seq
	TCTTTCAATTTTGAAGTACCATCTTGATTTTCTTATACTCCTCATTT Majority					
	8160	8170	8180	8190	8200	
8102	TCTTTCAATTTTGAAGTACCATCTTGATTTTCTTATACTCCTCATTT					2603_al2.seq
8102	TCTTTCAATTTTGAAGTACCATCTTGATTTTCTTATACTCCTCATTT					18rs21_al2.seq
8010	TCTTTCAATTTTGAAGTACCATCTTGATTTTCTTATACTCCTCATTT					515_al2.seq
7999	TCTTTCAATTTTGAAGTACCATCTTGATTTTCTTATACTCCTCATTT					cjb111_al2.seq
8052	TCTTTCAATTTTGAAGTACCATCTTGATTTTCTTATACTCCTCATTT					h36b_al2.seq
	ATAAACTTGTCTAAAACCCAGATATATCGATACCAAAATTA AAAATGTCAT Majority					
	8210	8220	8230	8240	8250	
8152	ATAAACTTGTCTAAAACCCAGATATATCGATACCAAAATTA AAAATGTCAT					2603_al2.seq
8152	ATAAACTTGTCTAAAACCCAGATATATCGATACCAAAATTA AAAATGTCAT					18rs21_al2.seq
8060	ATAAACTTGTCTAAAACCCAGATATATCGATACCAAAATTA AAAATGTCAT					515_al2.seq
8049	ATAAACTTGTCTAAAACCCAGATATATCGATACCAAAATTA AAAATGTCAT					cjb111_al2.seq
8102	ATAAACTTGTCTAAAACCCAGATATATCGATACCAAAATTA AAAATGTCAT					h36b_al2.seq
	AATTTTCTGTTTTAAACTATTTATATAAAAGTTTGCTTGCTGTTCCATGT Majority					
	8260	8270	8280	8290	8300	
8202	AATTTTCTGTTTTAAACTATTTATATAAAAGTTTGCTTGCTGTTCCATGT					2603_al2.seq
8202	AATTTTCTGTTTTAAACTATTTATATAAAAGTTTGCTTGCTGTTCCATGT					18rs21_al2.seq
8110	AATTTTCTGTTTTAAACTATTTATATAAAAGTTTGCTTGCTGTTCCATGT					515_al2.seq
8099	AATTTTCTGTTTTAAACTATTTATATAAAAGTTTGCTTGCTGTTCCATGT					cjb111_al2.seq
8152	AATTTTCTGTTTTAAACTATTTATATAAAAGTTTGCTTGCTGTTCCATGT					h36b_al2.seq
	TCTTTCACTGGTCCATTTTCGATAAAATTGTACCTTTAGGGTAATTAAGATT Majority					
	8310	8320	8330	8340	8350	
8252	TCTTTCACTGGTCCATTTTCGATAAAATTGTACCTTTAGGGTAATTAAGATT					2603_al2.seq
8252	TCTTTCACTGGTCCATTTTCGATAAAATTGTACCTTTAGGGTAATTAAGATT					18rs21_al2.seq
8160	TCTTTCACTGGTCCATTTTCGATAAAATTGTACCTTTAGGGTAATTAAGATT					515_al2.seq
8149	TCTTTCACTGGTCCATTTTCGATAAAATTGTACCTTTAGGGTAATTAAGATT					cjb111_al2.seq
8202	TCTTTCACTGGTCCATTTTCGATAAAATTGTACCTTTAGGGTAATTAAGATT					h36b_al2.seq
	TAAATCTAAATAATGAAGTTTTTGTAAAGTTTCCAGAGATTATCTGTGTTT Majority					
	8360	8370	8380	8390	8400	
8302	TAAATCTAAATAATGAAGTTTTTGTAAAGTTTCCAGAGATTATCTGTGTTT					2603_al2.seq
8302	TAAATCTAAATAATGAAGTTTTTGTAAAGTTTCCAGAGATTATCTGTGTTT					18rs21_al2.seq
8210	TAAATCTAAATAATGAAGTTTTTGTAAAGTTTCCAGAGATTATCTGTGTTT					515_al2.seq
8199	TAAATCTAAATAATGAAGTTTTTGTAAAGTTTCCAGAGATTATCTGTGTTT					cjb111_al2.seq
8252	TAAATCTAAATAATGAAGTTTTTGTAAAGTTTCCAGAGATTATCTGTGTTT					h36b_al2.seq
	GATAACTATCTAAGGGAAACAAAAGTAACCTCTCCCATTTTCCTTTTATA Majority					
	8410	8420	8430	8440	8450	
8352	GATAACTATCTAAGGGAAACAAAAGTAACCTCTCCCATTTTCCTTTTATA					2603_al2.seq
8352	GATAACTATCTAAGGGAAACAAAAGTAACCTCTCCCATTTTCCTTTTATA					18rs21_al2.seq
8260	GATAACTATCTAAGGGAAACAAAAGTAACCTCTCCCATTTTCCTTTTATA					515_al2.seq
8249	GATAACTATCTAAGGGAAACAAAAGTAACCTCTCCCATTTTCCTTTTATA					cjb111_al2.seq
8302	GATAACTATCTAAGGGAAACAAAAGTAACCTCTCCCATTTTCCTTTTATA					h36b_al2.seq
	TCCTCGGGCTTATCAGTAAGTAGAAAAATTACTTTTATTTAGATATCCATT Majority					
	8460	8470	8480	8490	8500	
8402	TCCTCGGGCTTATCAGTAAGTAGAAAAATTACTTTTATTTAGATATCCATT					2603_al2.seq
8402	TCCTCGGGCTTATCAGTAAGTAGAAAAATTACTTTTATTTAGATATCCATT					18rs21_al2.seq
8310	TCCTCGGGCTTATCAGTAAGTAGAAAAATTACTTTTATTTAGATATCCATT					515_al2.seq
8299	TCCTCGGGCTTATCAGTAAGTAGAAAAATTACTTTTATTTAGATATCCATT					cjb111_al2.seq
8352	TCCTCGGGCTTATCAGTAAGTAGAAAAATTACTTTTATTTAGATATCCATT					h36b_al2.seq
	TTTTTTCATTTGTTCAAATTTGGCTTTTCATATGATGCACCCAGTTTAAAAAT Majority					
	8510	8520	8530	8540	8550	
8452	TTTTTTCATTTGTTCAAATTTGGCTTTTCATATGATGCACCCAGTTTAAAAAT					2603_al2.seq
8452	TTTTTTCATTTGTTCAAATTTGGCTTTTCATATGATGCACCCAGTTTAAAAAT					18rs21_al2.seq
8360	TTTTTTCATTTGTTCAAATTTGGCTTTTCATATGATGCACCCAGTTTAAAAAT					515_al2.seq
8349	TTTTTTCATTTGTTCAAATTTGGCTTTTCATATGATGCACCCAGTTTAAAAAT					cjb111_al2.seq
8402	TTTTTTCATTTGTTCAAATTTGGCTTTTCATATGATGCACCCAGTTTAAAAAT					h36b_al2.seq

FIGURE 19R

75/487

T A T T A A T A G C A T A T G A T C T C G T A G G A A C A C C A T C A G T T A C A T G A A C A A T A Majority									
8560		8570		8580		8590		8600	
8502	T A T T A A T A G C A T A T G A T C T C G T A G G A A C A C C A T C A G T T A C A T G A A C A A T A 2603_al2.seq								
8502	T A T T A A T A G C A T A T G A T C T C G T A G G A A C A C C A T C A G T T A C A T G A A C A A T A 18rs21_al2.seq								
8410	T A T T A A T A G C A T A T G A T C T C G T A G G A A C A C C A T C A G T T A C A T G A A C A A T A 515_al2.seq								
8399	T A T T A A T A G C A T A T G A T C T C G T A G G A A C A C C A T C A G T T A C A T G A A C A A T A cjb111_al2.seq								
8452	T A T T A A T A G C A T A T G A T C T C G T A G G A A C A C C A T C A G T T A C A T G A A C A A T A h36b_al2.seq								
A T T T T T T G A C T A T T T C G A T T T A C T T G A C T C A A A A T A T C A T C T G C C T C C A T Majority									
8610		8620		8630		8640		8650	
8552	A T T T T T T G A C T A T T T C G A T T T A C T T G A C T C A A A A T A T C A T C T G C C T C C A T 2603_al2.seq								
8552	A T T T T T T G A C T A T T T C G A T T T A C T T G A C T C A A A A T A T C A T C T G C C T C C A T 18rs21_al2.seq								
8460	A T T T T T T G A C T A T T T C G A T T T A C T T G A C T C A A A A T A T C A T C T G C C T C C A T 515_al2.seq								
8449	A T T T T T T G A C T A T T T C G A T T T A C T T G A C T C A A A A T A T C A T C T G C C T C C A T cjb111_al2.seq								
8502	A T T T T T T G A C T A T T T C G A T T T A C T T G A C T C A A A A T A T C A T C T G C C T C C A T h36b_al2.seq								
G A A G G C T T T C A T A G T A A A T G T T T C T C C T A C T T T A C T A A G A T A G T A C T C C T Majority									
8660		8670		8680		8690		8700	
8602	G A A G G C T T T C A T A G T A A A T G T T T C T C C T A C T T T A C T A A G A T A G T A C T C C T 2603_al2.seq								
8602	G A A G G C T T T C A T A G T A A A T G T T T C T C C T A C T T T A C T A A G A T A G T A C T C C T 18rs21_al2.seq								
8510	G A A G G C T T T C A T A G T A A A T G T T T C T C C T A C T T T A C T A A G A T A G T A C T C C T 515_al2.seq								
8499	G A A G G C T T T C A T A G T A A A T G T T T C T C C T A C T T T A C T A A G A T A G T A C T C C T cjb111_al2.seq								
8552	G A A G G C T T T C A T A G T A A A T G T T T C T C C T A C T T T A C T A A G A T A G T A C T C C T h36b_al2.seq								
T T T G T T G C T C T G G A G T T A A T C C A T T G G T A G T A G A T C C C C A C T T A G C T T T A Majority									
8710		8720		8730		8740		8750	
8652	T T T G T T G C T C T G G A G T T A A T C C A T T G G T A G T A G A T C C C C A C T T A G C T T T A 2603_al2.seq								
8652	T T T G T T G C T C T G G A G T T A A T C C A T T G G T A G T A G A T C C C C A C T T A G C T T T A 18rs21_al2.seq								
8560	T T T G T T G C T C T G G A G T T A A T C C A T T G G T A G T A G A T C C C C A C T T A G C T T T A 515_al2.seq								
8549	T T T G T T G C T C T G G A G T T A A T C C A T T G G T A G T A G A T C C C C A C T T A G C T T T A cjb111_al2.seq								
8602	T T T G T T G C T C T G G A G T T A A T C C A T T G G T A G T A G A T C C C C A C T T A G C T T T A h36b_al2.seq								
G G A G C T T C T C T C G G A A T C C T T T T T A T A A T C T C T T C A G C A T T A T T T G T T A A Majority									
8760		8770		8780		8790		8800	
8702	G G A G C T T C T C T C G G A A T C C T T T T T A T A A T C T C T T C A G C A T T A T T T G T T A A 2603_al2.seq								
8702	G G A G C T T C T C T C G G A A T C C T T T T T A T A A T C T C T T C A G C A T T A T T T G T T A A 18rs21_al2.seq								
8610	G G A G C T T C T C T C G G A A T C C T T T T T A T A A T C T C T T C A G C A T T A T T T G T T A A 515_al2.seq								
8599	G G A G C T T C T C T C G G A A T C C T T T T T A T A A T C T C T T C A G C A T T A T T T G T T A A cjb111_al2.seq								
8652	G G A G C T T C T C T C G G A A T C C T T T T T A T A A T C T C T T C A G C A T T A T T T G T T A A h36b_al2.seq								
T T G T T T A T G A C T A T A A T T C T C T G T C T G A A T T G T G A A C T T A G T T T G A A G G C Majority									
8810		8820		8830		8840		8850	
8752	T T G T T T A T G A C T A T A A T T C T C T G T C T G A A T T G T G A A C T T A G T T T G A A G G C 2603_al2.seq								
8752	T T G T T T A T G A C T A T A A T T C T C T G T C T G A A T T G T G A A C T T A G T T T G A A G G C 18rs21_al2.seq								
8660	T T G T T T A T G A C T A T A A T T C T C T G T C T G A A T T G T G A A C T T A G T T T G A A G G C 515_al2.seq								
8649	T T G T T T A T G A C T A T A A T T C T C T G T C T G A A T T G T G A A C T T A G T T T G A A G G C cjb111_al2.seq								
8702	T T G T T T A T G A C T A T A A T T C T C T G T C T G A A T T G T G A A C T T A G T T T G A A G G C h36b_al2.seq								
C A T A A T A T T T A T C A T C T T C T T T A A A T C C T T T T A C G A C A T C T A C A C T C C T A Majority									
8860		8870		8880		8890		8900	
8802	C A T A A T A T T T A T C A T C T T C T T T A A A T C C T T T T A C G A C A T C T A C A C T C C T A 2603_al2.seq								
8802	C A T A A T A T T T A T C A T C T T C T T T A A A T C C T T T T A C G A C A T C T A C A C T C C T A 18rs21_al2.seq								
8710	C A T A A T A T T T A T C A T C T T C T T T A A A T C C T T T T A C G A C A T C T A C A C T C C T A 515_al2.seq								
8699	C A T A A T A T T T A T C A T C T T C T T T A A A T C C T T T T A C G A C A T C T A C A C T C C T A cjb111_al2.seq								
8752	C A T A A T A T T T A T C A T C T T C T T T A A A T C C T T T T A C G A C A T C T A C A C T C C T A h36b_al2.seq								
C C A T C A A A A A T A T C T G A A C C A T A G G T A A C T A A T G C A A C C C T A T T A T C A C T Majority									
8910		8920		8930		8940		8950	
8852	C C A T C A A A A A T A T C T G A A C C A T A G G T A A C T A A T G C A A C C C T A T T A T C A C T 2603_al2.seq								
8852	C C A T C A A A A A T A T C T G A A C C A T A G G T A A C T A A T G C A A C C C T A T T A T C A C T 18rs21_al2.seq								
8760	C C A T C A A A A A T A T C T G A A C C A T A G G T A A C T A A T G C A A C C C T A T T A T C A C T 515_al2.seq								
8749	C C A T C A A A A A T A T C T G A A C C A T A G G T A A C T A A T G C A A C C C T A T T A T C A C T cjb111_al2.seq								
8799	C C A T C A A A A A T A T C T G A A C C A T A G G T A A C T A A T G C A A C C C T A T T A T C A C T h36b_al2.seq								
G T T T G C T C C T A A A A T A T C T T T T A C T G C G G T C C C A A G A G C T T C G G C A G C T T Majority									
8960		8970		8980		8990		9000	
8902	G T T T G C T C C T A A A A T A T C T T T T A C T G C G G T C C C A A G A G C T T C G G C A G C T T 2603_al2.seq								
8902	G T T T G C T C C T A A A A T A T C T T T T A C T G C G G T C C C A A G A G C T T C G G C A G C T T 18rs21_al2.seq								
8810	G T T T G C T C C T A A A A T A T C T T T T A C T G C G G T C C C A A G A G C T T C G G C A G C T T 515_al2.seq								
8799	G T T T G C T C C T A A A A T A T C T T T T A C T G C G G T C C C A A G A G C T T C G G C A G C T T cjb111_al2.seq								
8849	G T T T G C T C C T A A A A T A T C T T T T A C T G C G G T C C C A A G A G C T T C G G C A G C T T h36b_al2.seq								

FIGURE 19S

FIGURE 19T

77/487

	T C T T T T T A T A A C C T T C G G G C C G C T G T T T C T T C T G A T A A A G T A T A A T C T C C A Majority																			
	9460				9470				9480				9490				9500			
9402	T C T T T T T A T A A C C T T C G G G C C G C T G T T T C T T C T G A T A A A G T A T A A T C T C C A 2603_a12.seq																			
9402	T C T T T T T A T A A C C T T C G G G C C G C T G T T T C T T C T G A T A A A G T A T A A T C T C C A 18rs21_a12.seq																			
9310	T C T T T T T A T A A C C T T C G G G C C G C T G T T T C T T C T G A T A A A G T A T A A T C T C C A 515_a12.seq																			
9299	T C T T T T T A T A A C C T T C G G G C C G C T G T T T C T T C T G A T A A A G T A T A A T C T C C A cjb111_a12.seq																			
9337	T C T T T T T T G T A T C C T T C G G G G G C G T G T T T C T T C T G A T A A A G T T A A T C T C C A h36b_a12.seq																			
	G G T A T G A G A T T A T C A A A A G T A G C T T C A C C T G T T A G C T C A G C A G T T A C T T T Majority																			
	9510				9520				9530				9540				9550			
9452	G G T A T G A G A T T A T C A A A A G T A G C T T C A C C T G T T A G C T C A G C A G T T A C T T T 2603_a12.seq																			
9452	G G T A T G A G A T T A T C A A A A G T A G C T T C A C C T G T T A G C T C A G C A G T T A C T T T 18rs21_a12.seq																			
9360	G G T A T G A G A T T A T C A A A A G T A G C T T C A C C T G T T A G C T C A G C A G T T A C T T T 515_a12.seq																			
9349	G G T A T G A G A T T A T C A A A A G T A G C T T C A C C T G T T A G C T C A G C A G T T A C T T T cjb111_a12.seq																			
9387	G G T T G C A G A T T A T C A A A A G T A G C T T C A C C T G T T A G C T C A G C A G T T A C T T T h36b_a12.seq																			
	T T C T A T T T T A C T T T C T G G A T G A G C A G T A G T T T T T A A A A C A A A G G T A G C T T Majority																			
	9560				9570				9580				9590				9600			
9502	T T C T A T T T T A C T T T C T G G A T G A G C A G T A G T T T T T A A A A C A A A G G T A G C T T 2603_a12.seq																			
9502	T T C T A T T T T A C T T T C T G G A T G A G C A G T A G T T T T T A A A A C A A A G G T A G C T T 18rs21_a12.seq																			
9410	T T C T A T T T T A C T T T C T G G A T G A G C A G T A G T T T T T A A A A C A A A G G T A G C T T 515_a12.seq																			
9399	T T C T A T T T T A C T T T C T G G A T G A G C A G T A G T T T T T A A A A C A A A G G T A G C T T cjb111_a12.seq																			
9437	T T C T A T T T T T C T T T T C T G A G T G T G A G T A G T T T T T A A A A C A A A G G T A G C T T h36b_a12.seq																			
	T T G A A A G T G C T T T G T T C T G G T C A T C T G T C T T T T T A A C A A C T A A C T T T C C T Majority																			
	9610				9620				9630				9640				9650			
9552	T T G A A A G T G C T T T G T T C T G G T C A T C T G T C T T T T T A A C A A C T A A C T T T C C T 2603_a12.seq																			
9552	T T G A A A G T G C T T T G T T C T G G T C A T C T G T C T T T T T A A C A A C T A A C T T T C C T 18rs21_a12.seq																			
9460	T T G A A A G T G C T T T G T T C T G G T C A T C T G T C T T T T T A A C A A C T A A C T T T C C T 515_a12.seq																			
9449	T T G A A A G T G C T T T G T T C T G G T C A T C T G T C T T T T T A A C A A C T A A C T T T C C T cjb111_a12.seq																			
9487	T T G A A A G T G C T T T G T T C T G G T C A T C T G T C T T T T T A A C A A C T A A C T T T C C T h36b_a12.seq																			
	T T A G C A C C A T T T T C C G G T A C G G T A C T T T C C C C T A A A A C A T T G G T A T T A A G Majority																			
	9660				9670				9680				9690				9700			
9602	T T A G C A C C A T T T T C C G G T A C G G T A C T T T C C C C T A A A A C A T T G G T A T T A A G 2603_a12.seq																			
9602	T T A G C A C C A T T T T C C G G T A C G G T A C T T T C C C C T A A A A C A T T G G T A T T A A G 18rs21_a12.seq																			
9510	T T A G C A C C A T T T T C C G G T A C G G T A C T T T C C C C T A A A A C A T T G G T A T T A A G 515_a12.seq																			
9499	T T A G C A C C A T T T T C C G G T A C G G T A C T T T C C C C T A A A A C A T T G G T A T T A A G cjb111_a12.seq																			
9537	T T A G C A C C A T T T T C C G G T A C G G T A C T T T C C C C T A A A A C A T T G G T A T T A A G h36b_a12.seq																			
	C G G T A T T T G C G A C A A A C A A A A A G A C T T A A C G T C A A T A T T T T A G A A A A T T Majority																			
	9710				9720				9730				9740				9750			
9652	C G G T A T T T G C G A C A A A C A A A A A G A C T T A A C G T C A A T A T T T T A G A A A A T T 2603_a12.seq																			
9652	C G G T A T T T G C G A C A A A C A A A A A G A C T T A A C G T C A A T A T T T T A G A A A A T T 18rs21_a12.seq																			
9560	C G G T A T T T G C G A C A A A C A A A A A G A C T T A A C G T C A A T A T T T T A G A A A A T T 515_a12.seq																			
9549	C G G T A T T T G C G A C A A A C A A A A A G A C T T A A C G T C A A T A T T T T A G A A A A T T cjb111_a12.seq																			
9587	C G G T A T T T G C G A C A A A C A A A A A G A C T T A A C G T C A A T A T T T T A G A A A A T T h36b_a12.seq																			
	T T T G G T A T T T T C T C A T T T T A C A A C T C C T A T T G T G C C G A A A T G T C G T T T C T Majority																			
	9760				9770				9780				9790				9800			
9702	T T T G G T A T T T T C T C A T T T T A C A A C T C C T A T T G T G C C G A A A T G T C G T T T C T 2603_a12.seq																			
9702	T T T G G T A T T T T C T C A T T T T A C A A C T C C T A T T G T G C C G A A A T G T C G T T T C T 18rs21_a12.seq																			
9610	T T T G G T A T T T T C T C A T T T T A C A A C T C C T A T T G T G C C G A A A T G T C G T T T C T 515_a12.seq																			
9599	T T T G G T A T T T T C T C A T T T T A C A A C T C C T A T T G T G C C G A A A T G T C G T T T C T cjb111_a12.seq																			
9637	T T T G G T A T T T T C T C A T T T T A C A A C T C C T A T T G T G C C G A A A T G T C G T T T C T h36b_a12.seq																			
	A A A T C T A A G A T C A G A T A C A G A A T A T C C T A G A A T A T A C A A A C T A T C A C T T A Majority																			
	9810				9820				9830				9840				9850			
9752	A A A T C T A A G A T C A G A T A C A G A A T A T C C T A G A A T A T A C A A A C T A T C A C T T A 2603_a12.seq																			
9752	A A A T C T A A G A T C A G A T A C A G A A T A T C C T A G A A T A T A C A A A C T A T C A C T T A 18rs21_a12.seq																			
9660	A A A T C T A A G A T C A G A T A C A G A A T A T C C T A G A A T A T A A A A A C T A T C A C T T A 515_a12.seq																			
9649	A A A T C T A A G A T C A G A T A C A G A A T A T C C T A G A A T A T A C A A A C T A T C A C T T A cjb111_a12.seq																			
9687	A A A T C T A A G A T C A G A T A C A G A A T A T C C T A G A A T A T A C A A A C T A T C A C T T A h36b_a12.seq																			
	T T A T G A T A T C A A T A A T T T C T T A T T A T A A G C T A T G C A A T T T T A A T G T T T T Majority																			
	9860				9870				9880				9890				9900			
9802	T T A T G A T A T C A A T A A T T T C T T A T T A T A A G G T A T G C A A T T T T A A T G T T T T 2603_a12.seq																			
9802	T T A T G A T A T C A A T A A T T T C T T A T T A T A A G G T A T G C A A T T T T A A T G T T T T 18rs21_a12.seq																			
9710	T T A T G A T A T C A A T A A T T T C T T A T T A T A A G G T A T G C A A T T T T A A G T T T T T 515_a12.seq																			
9699	T T A T G A T A T C A A T A A T T T C T T A T T A T A A G C T A T G C A A T T T T A A T G T T T T cjb111_a12.seq																			
9737	T T A T G A T A T C A A T A A T T T C T T A T T A T A A G G T A T G C A A T T T T A A G T T T T T h36b_a12.seq																			

78/487

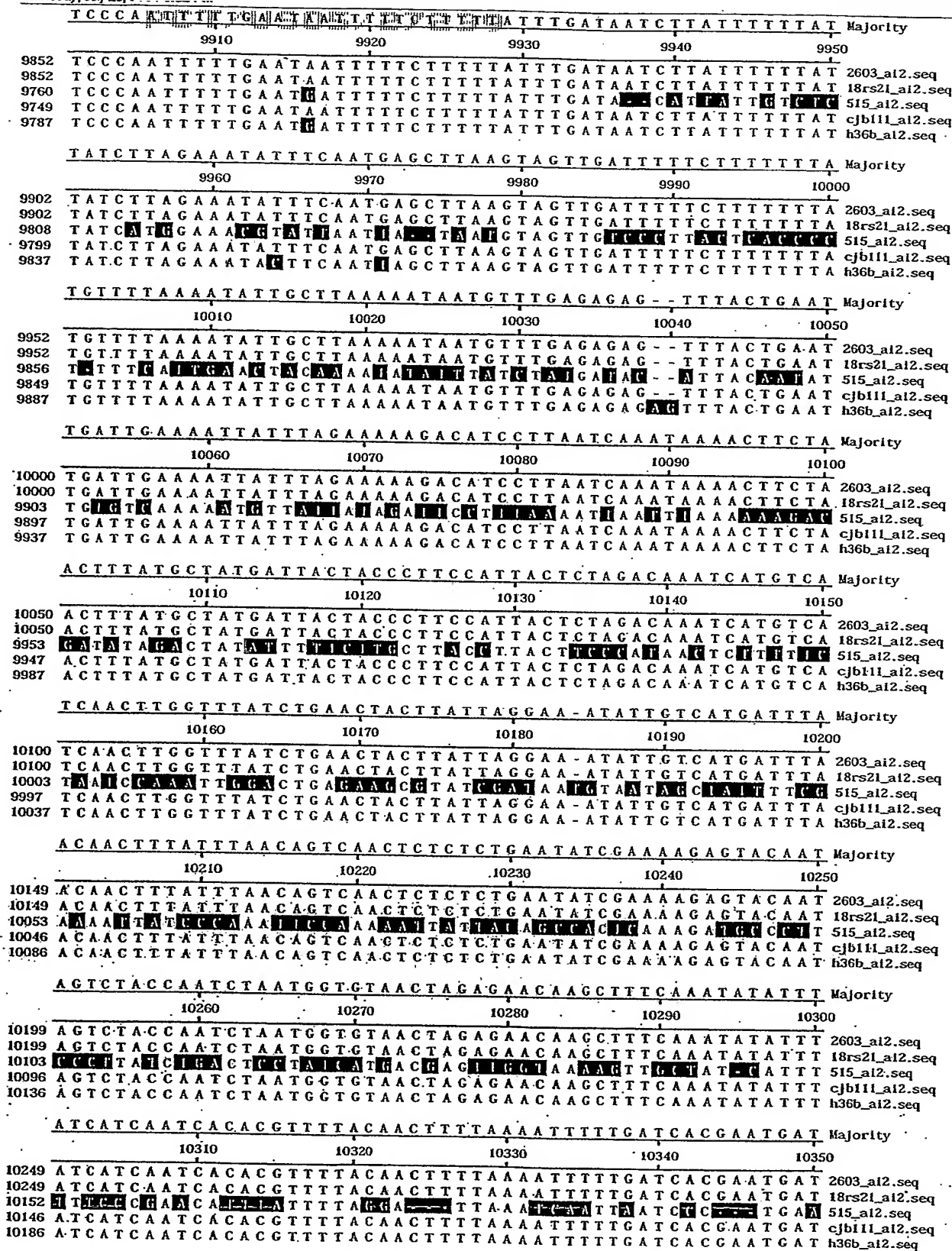


FIGURE 19V

FIGURE 19W

FIGURE 19X

8/1487

ATACAGAGTATAGCCCTACTAACTACAGAGCTTTTATTAATAACCAAATTTTAA Majority
11260 11270 11280 11290 11300

11174 ATACAGAGTATAGCCCTACTAACTACAGAGCTTTTATTAATAACCAAATTTTAA 2603_a12.seq
11175 ATACAGAGTATAGCCCTACTAACTACAGAGCTTTTATTAATAACCAAATTTTAA 18rs21_a12.seq
11086 ATACAGAGTATAGCCCTACTAACTACAGAGCTTTTATTAATAACCAAATTTTAA 515_a12.seq
11072 ATACAGAGTATAGCCCTACTAACTACAGAGCTTTTATTAATAACCAAATTTTAA cjb111_a12.seq
11112 ATACAGAGTATAGCCCTACTAACTACAGAGCTTTTATTAATAACCAAATTTTAA h36b_a12.seq

ACAGAAATGCTTTATTACAGAGATTTTCTTCTAAAAAGATTTCATTTCCACCC Majority
11310 11320 11330 11340 11350

11222 ACAGAAATGCTTTATTACAGAGATTTTCTTCTAAAAAGATTTCATTTCCACCC 2603_a12.seq
11223 ACAGAAATGCTTTATTACAGAGATTTTCTTCTAAAAAGATTTCATTTCCACCC 18rs21_a12.seq
11136 ATAGGAATGATTTTCTTCTAAAAAGATTTCATTTCCACCC 515_a12.seq
11120 ACAGAAATGCTTTATTACAGAGATTTTCTTCTAAAAAGATTTCATTTCCACCC cjb111_a12.seq
11160 ACAGAAATGCTTTATTACAGAGATTTTCTTCTAAAAAGATTTCATTTCCACCC h36b_a12.seq

TTACTATTTTATTAACCTGATGATCTTTTCCAATATTACTAATCTTAATCCAG Majority
11360 11370 11380 11390 11400

11272 TTACTATTTTATTAACCTGATGATCTTTTCCAATATTACTAATCTTAATCCAG 2603_a12.seq
11273 TTACTATTTTATTAACCTGATGATCTTTTCCAATATTACTAATCTTAATCCAG 18rs21_a12.seq
11181 TTACTATTTTATTAACCTGATGATCTTTTCCAATATTACTAATCTTAATCCAG 515_a12.seq
11170 TTACTATTTTATTAACCTGATGATCTTTTCCAATATTACTAATCTTAATCCAG cjb111_a12.seq
11210 TTACTATTTTATTAACCTGATGATCTTTTCCAATATTACTAATCTTAATCCAG h36b_a12.seq

ACATTATTTATTAACCAATAAAAAAGCTTTTCTCCCTTTATCAAAACATGAGATT Majority
11410 11420 11430 11440 11450

11322 ACATTATTTATTAACCAATAAAAAAGCTTTTCTCCCTTTATCAAAACATGAGATT 2603_a12.seq
11323 ACATTATTTATTAACCAATAAAAAAGCTTTTCTCCCTTTATCAAAACATGAGATT 18rs21_a12.seq
11227 ATAGGAATGATTTTCTTCTAAAAAGATTTCATTTCCACCC 515_a12.seq
11220 ACATTATTTATTAACCAATAAAAAAGCTTTTCTCCCTTTATCAAAACATGAGATT cjb111_a12.seq
11260 ACATTATTTATTAACCAATAAAAAAGCTTTTCTCCCTTTATCAAAACATGAGATT h36b_a12.seq

TCTTCAGAGAGTTTAAATTACATATATTGATCTGATTAAATACTTCAGACCA Majority
11460 11470 11480 11490 11500

11372 TCTTCAGAGAGTTTAAATTACATATATTGATCTGATTAAATACTTCAGACCA 2603_a12.seq
11373 TCTTCAGAGAGTTTAAATTACATATATTGATCTGATTAAATACTTCAGACCA 18rs21_a12.seq
11269 TCTTCAGAGAGTTTAAATTACATATATTGATCTGATTAAATACTTCAGACCA 515_a12.seq
11270 TCTTCAGAGAGTTTAAATTACATATATTGATCTGATTAAATACTTCAGACCA cjb111_a12.seq
11310 TCTTCAGAGAGTTTAAATTACATATATTGATCTGATTAAATACTTCAGACCA h36b_a12.seq

AATCAATCAAAATCCAAAAAATTATTTTCATCAATACAGGAAGAAAAATATT Majority
11510 11520 11530 11540 11550

11422 AATCAATCAAAATCCAAAAAATTATTTTCATCAATACAGGAAGAAAAATATT 2603_a12.seq
11423 AATCAATCAAAATCCAAAAAATTATTTTCATCAATACAGGAAGAAAAATATT 18rs21_a12.seq
11319 ATAGGAATGATTTTCTTCTAAAAAGATTTCATTTCCACCC 515_a12.seq
11320 AATCAATCAAAATCCAAAAAATTATTTTCATCAATACAGGAAGAAAAATATT cjb111_a12.seq
11360 AATCAATCAAAATCCAAAAAATTATTTTCATCAATACAGGAAGAAAAATATT h36b_a12.seq

GCAAACTTTTTTGCAAAAACTAATGAAATAACTAATCGTAGCTCCT - - - Majority
11560 11570 11580 11590 11600

11472 GCAAACTTTTTTGCAAAAACTAATGAAATAACTAATCGTAGCTCCT - - - 2603_a12.seq
11473 GCAAACTTTTTTGCAAAAACTAATGAAATAACTAATCGTAGCTCCT - - - 18rs21_a12.seq
11369 ATAGGAATGATTTTCTTCTAAAAAGATTTCATTTCCACCC 515_a12.seq
11370 GCAAACTTTTTTGCAAAAACTAATGAAATAACTAATCGTAGCTCCT - - - cjb111_a12.seq
11410 GCAAACTTTTTTGCAAAAACTAATGAAATAACTAATCGTAGCTCCT - - - h36b_a12.seq

ATAACTCTTAAAAAATTAAACATTAAAAAGCT - AGAGCATTGTGTAATGCTC Majority
11610 11620 11630 11640 11650

11517 ATAACTCTTAAAAAATTAAACATTAAAAAGCT - AGAGCATTGTGTAATGCTC 2603_a12.seq
11518 ATAACTCTTAAAAAATTAAACATTAAAAAGCT - AGAGCATTGTGTAATGCTC 18rs21_a12.seq
11419 ATAAGGAATGATTTTCTTCTAAAAAGATTTCATTTCCACCC 515_a12.seq
11415 ATAACTCTTAAAAAATTAAACATTAAAAAGCT - AGAGCATTGTGTAATGCTC cjb111_a12.seq
11455 ATAACTCTTAAAAAATTAAACATTAAAAAGCT - AGAGCATTGTGTAATGCTC h36b_a12.seq

TAGCTTTTTTAAATGTTAATTTTTTTTGAATAATATAATCCAACCTTTTCAACT Majority
11660 11670 11680 11690 11700

11566 TAGCTTTTTTAAATGTTAATTTTTTTTGAATAATATAATCCAACCTTTTCAACT 2603_a12.seq
11567 TAGCTTTTTTAAATGTTAATTTTTTTTGAATAATATAATCCAACCTTTTCAACT 18rs21_a12.seq
11466 ATAGGAATGATTTTCTTCTAAAAAGATTTCATTTCCACCC 515_a12.seq
11464 TAGCTTTTTTAAATGTTAATTTTTTTTGAATAATATAATCCAACCTTTTCAACT cjb111_a12.seq
11504 TAGCTTTTTTAAATGTTAATTTTTTTTGAATAATATAATCCAACCTTTTCAACT h36b_a12.seq

FIGURE 19Y

82/487

G T T T T T T C C A T G T G A A A T G T C T T T T A G C A A T A T T C T G T T G Majority										
11710		11720		11730		11740		11750		
11616	G T T T T T T C C C A T G T G A A A T G T T C T T T A A T T C T T T T A G C A A T A T T C T G T T G									2603_a12.seq
11617	G T T T T T T C C C A T G T G A A A T G T T C T T T A A T T C T T T T A G C A A T A T T C T G T T G									18rs21_a12.seq
11514	G A A A A A T G A C T G A C C C A T G T C T A A A A A A A T A C T A G C A A A T G G T G T T									515_a12.seq
11514	G T T T T T T C C C A T G T G A A A T G T T C T T T A A T T C T T T T A G C A A T A T T C T G T T G									cjb111_a12.seq
11554	G T T T T T T C C C A T G T G A A A T G T T C T T T A A T T C T T T T A G C A A T A T T C T G T T G									h36b_a12.seq
T A G T T T C T C T C T T A A T G C C T T A T C T T T T A C T A A T A A A T C A A G A G A T T C A T Majority										
11760		11770		11780		11790		11800		
11666	T A G T T T C T C T C T T A A T G C C T T A T C T T T T A C T A A T A A A T C A A G A G A T T C A T									2603_a12.seq
11667	T A G T T T C T C T C T T A A T G C C T T A T C T T T T A C T A A T A A A T C A A G A G A T T C A T									18rs21_a12.seq
11564	C T T T G C A A A A G A T A G C C C T T C T A A C G T T G T T A A C T G A A A A C C T C T A T T A C									515_a12.seq
11564	T A G T T T C T C T C T T A A T G C C T T A T C T T T T A C T A A T A A A T C A A G A G A T T C A T									cjb111_a12.seq
11604	T A G T T T C T C T C T T A A T G C C T T A T C T T T T A C T A A T A A A T C A A G A G A T T C A T									h36b_a12.seq
G G A G T - - - G A C T G A G T A T T T T C T T C C A T G A T G A T T C C T A A C T C A G G G C T A Majority										
11810		11820		11830		11840		11850		
11716	G G A G T - - - G A C T G A G T A T T T T C T T C C A T G A T G A T T C C T A A C T C A G G G C T A									2603_a12.seq
11717	G G A G T - - - G A C T G A G T A T T T T C T T C C A T G A T G A T T C C T A A C T C A G G G C T A									18rs21_a12.seq
11614	A A A A C A C A A A A T G A A T A T T T T C A T C T G A A A C A T A T T G A C A G G G C T T C									515_a12.seq
11614	G G A G T - - - G A C T G A G T A T T T T C T T C C A T G A T G A T T C C T A A C T C A G G G C T A									cjb111_a12.seq
11654	G G A G T - - - G A C T G A G T A T T T T C T T C C A T G A T G A T T C C T A A C T C A G G G C T A									h36b_a12.seq
T - - - - - C A A T A A C T T C A A C T G T T C C A C C G C G A T C T G T T G C A A T A A T A G C Majority										
11860		11870		11880		11890		11900		
11763	T - - - - - C A A T A A C T T C A A C T G T T C C A C C G C G A T C T G T T G C A A T A A T A G C									2603_a12.seq
11764	T - - - - - C A A T A A C T T C A A C T G T T C C A C C G C G A T C T G T T G C A A T A A T A G C									18rs21_a12.seq
11664	T A T T C T G A T T T T T G T T A A I A A T A G C A I A G C C T T G C T T G A C A A T A T T T T									515_a12.seq
11661	T - - - - - C A A T A A C T T C A A C T G T T C C A C C G C G A T C T G T T G C A A T A A T A G C									cjb111_a12.seq
11701	T - - - - - C A A T A A C T T C A A C T G T T C C A C C G C G A T C T G T T G C A A T A A T A G C									h36b_a12.seq
A C T T G A A A G T A G A C C A G C T T C T A A A - - - - A T A G A G G T T G G T A A T C C C T C Majority										
11910		11920		11930		11940		11950		
11807	A C T T G A A A G T A G A C C A G C T T C T A A A - - - - A T A G A G G T T G G T A A T C C C T C									2603_a12.seq
11808	A C T T G A A A G T A G A C C A G C T T C T A A A - - - - A T A G A G G T T G G T A A T C C C T C									18rs21_a12.seq
11714	G G T T G G G I A A A C G G I A A A T T T T T T I A C C C I I G T T T C A T C I A T A A T C G G T									515_a12.seq
11705	A C T T G A A A G T A G A C C A G C T T C T A A A - - - - A T A G A G G T T G G T A A T C C C T C									cjb111_a12.seq
11745	A C T T G A A A G T A G A C C A G C T T C T A A A - - - - A T A G A G G T T G G T A A T C C C T C									h36b_a12.seq
T G G A T A C A T T G A A G G G T A A A C A A A G A T A T C A G T C T G T - G C C A T T A A A G A C Majority										
11960		11970		11980		11990		12000		
11852	T G G A T									

FIGURE 19Z

FIGURE 19AA

FIGURE 19AB

FIGURE 19AC

FIGURE 19AD

87/487

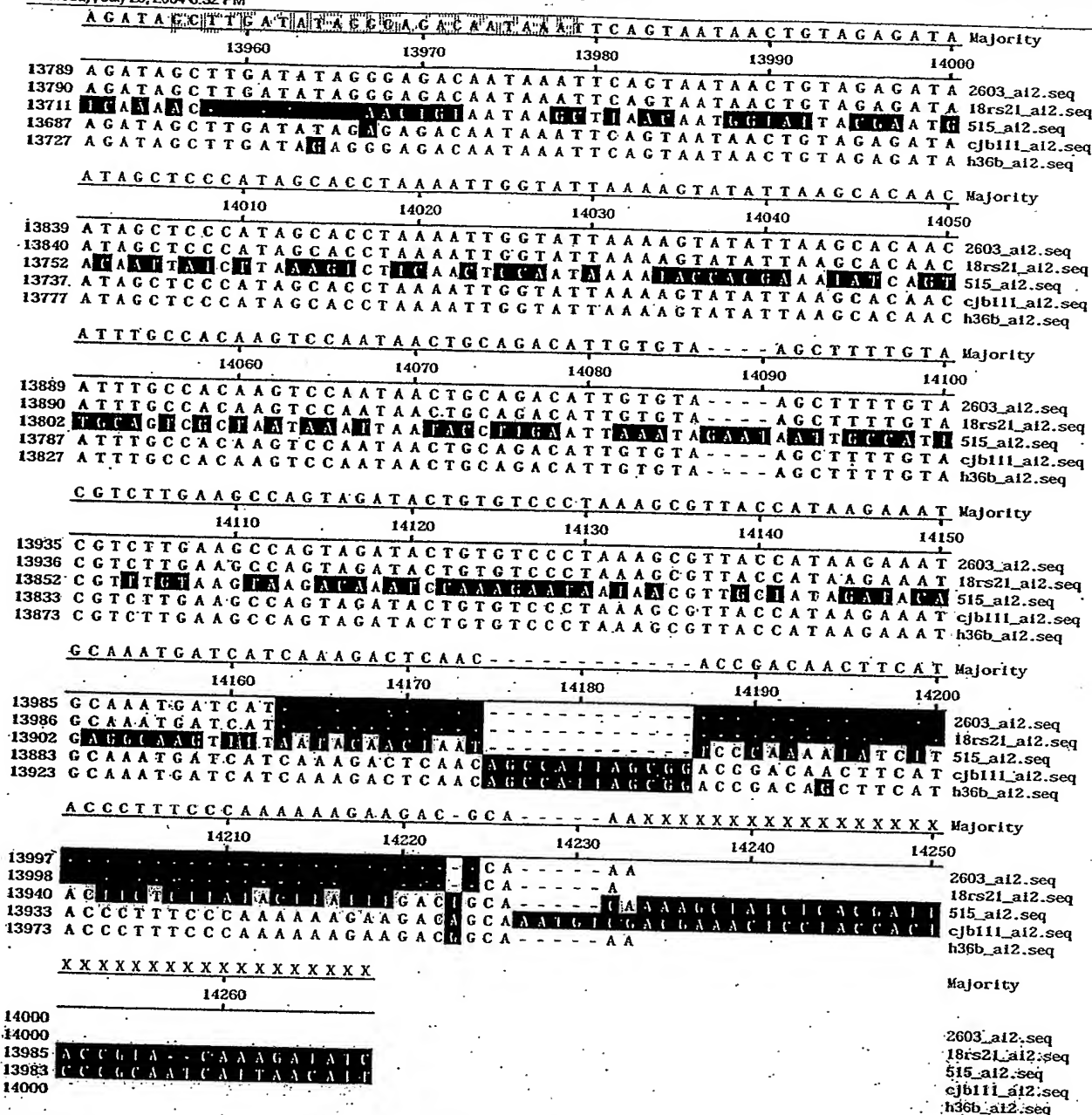


FIGURE 19AE

88/487

		G T T G C T T C A T A A G T T G T C T G A A G C C T A A T T C T A A A G T C A C A T G C A T T C T	Majority
		10 20 30 40 50	
1		G	
1		G T T G C T T C A T A A G T T G T C T G A A G C C T A A T T C T A A A G T C A C A T G C A T T C T	2603_ai2.seq
		T T C A G A A A G T T C A G C G A G A T A G T A T A T A G T T T C A T C A G G T A A G C A A T C C G	nem316_ai2.seq
		60 70 80 90 100	Majority
2		T T C A G A A A G T T C A G C G A G A T A G T A T A T A G T T T C A T C A G G T A A G C A A T C C G	2603_ai2.seq
51		T T C A G A A A G T T C A G C G A G A T A G T A T A T A G T T T C A T C A G G T A A G C A A T C C G	nem316_ai2.seq
		G C C T T G T T C C G A T C T T G A T T C C G A T A A C T C C T G G C T C A T T A A T A G C C T G T	Majority
		110 120 130 140 150	
2		G C C T T G T T C C G A T C T T G A T T C C G A T A A C T C C T G G C T C A T T A A T A G C C T G T	2603_ai2.seq
101		G C C T T G T T C C G A T C T T G A T T C C G A T A A C T C C T G G C T C A T T A A T A G C C T G T	nem316_ai2.seq
		T C G T A A C G C T G T T T A A T T A T C T C T A A C T T A G C A T G G G T A T T G G T A A A A T T	Majority
		160 170 180 190 200	
52		T C G T A A C G C T G T T T A A T T A T C T C T A A C T T A G C A T G G G T A T T G G T A A A A T T	2603_ai2.seq
151		T C G T A A C G C T G T T T A A T T A T C T C T A A C T T A G C A T G G G T A T T G G T A A A A T T	nem316_ai2.seq
		T T C A A A A T A G A C T A A G T A T T T A T T A A C C T C G G G C C A C T T T C T A T G C A T G A	Majority
		210 220 230 240 250	
102		T T G A A A A T A G A C T A A G T A T T T A T T A A C C T C A G G C C A C T T T C T A T G C A T G A	2603_ai2.seq
201		T T G A A A A T A G A C T A A G T A T T T A T T A A C C T C G G G C C A C T T T C T A T G C A T G A	nem316_ai2.seq
		A A T C A A T T T C T T T A T A G A A T T G T T C A C G A A T A G G A G C T T C T G G A G C A A C T	Majority
		260 270 280 290 300	
152		A A T C A A T T T C T T T A T A G A A T T G T T C A C G A A T A G G A G C T T C T G G A G C A A C T	2603_ai2.seq
251		A A T C A A T T T C T T T A T A G A A T T G T T C A C G A A T A G G A G C T T C T G G A G C A A C T	nem316_ai2.seq
		A T A G C A T C C C C T G A A C C A G A A A C T G T G C A A A A A G T G C A T C C T C C T C T A G C	Majority
		310 320 330 340 350	
202		A T A G C A T C C C C T G A A C C A G A A A C T G T G C A A A A A G T G C A T C C T C C T C T A G C	2603_ai2.seq
301		A T A G C A T C C C C T G A A C C A G A A A C T G T G C A A A A A G T G C A T C C T C C T C T A G C	nem316_ai2.seq
		A A C T G T T C C G T C T C T G T T A G G A C A G T C A A A A C C A G C A T C T A T A G G T A A T T	Majority
		360 370 380 390 400	
252		A A C T G T T C C G T C T C T G T T A G G A C A G T C A A A A C C A G C A T C T A T A G G T A A T T	2603_ai2.seq
351		A A C T G T T C C G T C T C T G T T A G G A C A G T C A A A A C C A G C A T C T A T A G G T A A T T	nem316_ai2.seq
		T A A A T A T T T T T C T C C A A A G A G T T C T C G A T A A T A A T C A T T A A T C G C A C G A	Majority
		410 420 430 440 450	
302		T A A A T A T T T T T C T C C A A A G A G T T C T C G A T A A T A A T C A T T A A T C G C A C G A	2603_ai2.seq
401		T A A A T A T T T T T C T C C A A A G A G T T C T C G A T A A T A A T C A T T A A T C G C A C G A	nem316_ai2.seq
		T A A C G T T T T T T C A T A G G A T A A T T G T A T C A C A A T T T T A A C T A A A A T A A C C T	Majority
		460 470 480 490 500	
352		T A A C G T T T T T T C A T A G G A T A A T T G T A T C A C A A T T T T A A C T A A A A T A A C C T	2603_ai2.seq
451		T A A C G T T T T T T C A T A G G A T A A T T G T A T C A C A A T T T T A A C T A A A A T A A C C T	nem316_ai2.seq
		C A C T A C T A C A A T A A A A C T T A A A A A G A T T G G A A C G T C A G T T A G T C C G A A T C	Majority
		510 520 530 540 550	
402		C A C T A C T A C A A T A A A A C T T A A A A A G A T T G G A A C G T C A G T T A G T C C G A A T C	2603_ai2.seq
501		C A C T A C T A C A A T A A A A C T T A A A A A G A T T G G A A C G T C A G T T A G T C C G A A T C	nem316_ai2.seq
		T T T T A T T T A C T T C A C T T T C T T T A A C C A A T C C T T G G C T A A A A A G A T A T A C G	Majority
		560 570 580 590 600	
452		T T T T A T T T A C T T C A C T T T C T T T A A C C A A T C C T T G G C T A A A A A G A T A T A C G	2603_ai2.seq
551		T T T T A T T T A C T T C A C T T T C T T T A A C C A A T C C T T G G C T A A A A A G A T A T A C G	nem316_ai2.seq
		C A G T T A G A T T C A A A A T A C C A T A A G C A A G T A T A A A A C C A G C T A A A A C A T C T	Majority
		610 620 630 640 650	
502		C A G T T A G A T T C A A A A T A C C A T A A G C A A G T A T A A A A C C A G C T A A A A C A T C T	2603_ai2.seq
601		C A G T T A G A T T C A A A A T A C C A T A A G C A A G T A T A A A A C C A G C T A A A A C A T C T	nem316_ai2.seq

Figure 20

89/487

	GTCTGGGAAATGCAACCCCTAGGTAATAACGAGATAACCCCAATTAAAAAAAT	Majority
552	GTCTGGGAAATGCAACCCCTAGGTAATAACGAGATAACCCCAATTAAAAAAAT	2603_a12.seq
651	GTCTGGGAAATGCAACCCCTAGGTAATAACGAGATAACCCCAATTAAAAAAAT	nem316_a12.seq
	GAGCAAACCCAAAGTACCTTTGGCACAACAGTTTCCATATACTCTTAGGCCA	Majority
602	GAGCAAACCCAAAGTACCTTTGGCACAACAGTTTCCATATACTCTTAGGCCA	2603_a12.seq
701	GAGCAAACCCAAAGTACCTTTGGCACAACAGTTTCCATATACTCTTAGGCCA	nem316_a12.seq
	TATAGTACTGCAATAAAAAATAATAACTCCCAAATATCATAAATGTTCCC	Majority
652	TATAGTACTGCAATAAAAAATAATAACTCCCAAATATCATAAATGTTCCC	2603_a12.seq
751	TATAGTACTGCAATAAAAAATAATAACTCCCAAATATCATAAATGTTCCC	nem316_a12.seq
	ATCGAGTGGCCACTGGGAAACGAATAGCCACCTGCAAATACTAAATGGGT	Majority
702	ATCGAGTGGCCACTGGGAAACGAATAGCCACCTGCAAATACTAAATGGGT	2603_a12.seq
801	ATCGAGTGGCCACTGGGAAACGAATAGCCACCTGCAAATACTAAATGGGT	nem316_a12.seq
	TAAAGTTGGTCTTACTCTTTGAAAAATAAGTTTAAAGAAAGTATACATA	Majority
752	TAAAGTTGGTCTTACTCTTTGAAAAATAAGTTTAAAGAAAGTATACATA	2603_a12.seq
851	TAAAGTTGGTCTTACTCTTTGAAAAATAAGTTTAAAGAAAGTATACATA	nem316_a12.seq
	TACCAGAGATAATAGCATTCTTACTGCGATAAATCTAGCTTGAGGATACCAC	Majority
802	TACCAGAGATAATAGCATTCTTACTGCGATAAATCTAGCTTGAGGATACCAC	2603_a12.seq
901	TACCAGAGATAATAGCATTCTTACTGCGATAAATCTAGCTTGAGGATACCAC	nem316_a12.seq
	TTCTTAAGGTAACAGAAAGTGACGCTCATAATCGCAATAGCTATCTGGCT	Majority
852	TTCTTAAGGTAACAGAAAGTGACGCTCATAATCGCAATAGCTATCTGGCT	2603_a12.seq
951	TTCTTAAGGTAACAGAAAGTGACGCTCATAATCGCAATAGCTATCTGGCT	nem316_a12.seq
	TACAGTATTACCAATCACAGTGATTAACCTTGAAAATCTTGTAGAAAGAT	Majority
902	TACAGTATTACCAATCACAGTGATTAACCTTGAAAATCTTGTAGAAAGAT	2603_a12.seq
1001	TACAGTATTACCAATCACAGTGATTAACCTTGAAAATCTTGTAGAAAGAT	nem316_a12.seq
	TTGGCAACTGTCTCTTAACACTTTTCTTGAATGTTTTGGTCAAATGCAATT	Majority
952	TTGGCAACTGTCTCTTAACACTTTTCTTGAATGTTTTGGTCAAATGCAATT	2603_a12.seq
1051	TTGGCAACTGTCTCTTAACACTTTTCTTGAATGTTTTGGTCAAATGCAATT	nem316_a12.seq
	ACAGTGTGCGGGCCCAATATTTGATGACCAATCCTAAACTGAAAAATAAGAT	Majority
1002	ACAGTGTGCGGGCCCAATATTTGATGACCAATCCTAAACTGAAAAATAAGAT	2603_a12.seq
1101	ACAGTGTGCGGGCCCAATATTTGATGACCAATCCTAAACTGAAAAATAAGAT	nem316_a12.seq
	AATAGCAATAAATGCTTGAATAAGTTTACTATTTTGACGAGATAACATTA	Majority
1052	AATAGCAATAAATGCTTGAATAAGTTTACTATTTTGACGAGATAACATTA	2603_a12.seq
1151	AATAGCAATAAATGCTTGAATAAGTTTACTATTTTGACGAGATAACATTA	nem316_a12.seq
	GTCTTTTTTATATCTTTCTAATATTGGCAAACAAGCCACGTAAGTTAGATA	Majority
1102	GTCTTTTTTATATCTTTCTAATATTGGCAAACAAGCCACGTAAGTTAGATA	2603_a12.seq
1201	GTCTTTTTTATATCTTTCTAATATTGGCAAACAAGCCACGTAAGTTAGATA	nem316_a12.seq
	GAAAAACAATCGAAATTAAAAATTCCCTCAACGATATTAAATGGAATAACCA	Majority
1152	GAAAAACAATCGAAATTAAAAATTCCCTCAACGATATTAAATGGAATAACCA	2603_a12.seq
1251	GAAAAACAATCGAAATTAAAAATTCCCTCAACGATATTAAATGGAATAACCA	nem316_a12.seq

FIGURE 20A

90/487

	TTGTTAAAGGCTAATTGCTACACCAATAAATGTTCTGATATCAAAGTTA	Majority
	1310 1320 1330 1340 1350	
1202	TTGTTAAAGGCTAATTGCTACACCAATAAATGTTCTGATATCAAAGTTA	2603_a12.seq
1301	TTGTTAAAGGCTAATTGCTACACCAATAAATGTTCTGATATCAAAGTTA	nem316_a12.seq
	GCAAAATATAGCATACAAAGGAATCGCAAAAGACATAGTTGAGAGCTACCAT	Majority
	1360 1370 1380 1390 1400	
1252	GCAAAATATAGCATACAAAGGAATCGCAAAAGACATAGTTGAGAGCTACCAT	2603_a12.seq
1351	GCAAAATATAGCATACAAAGGAATCGCAAAAGACATAGTTGAGAGCTACCAT	nem316_a12.seq
	AGATACGGTCAAGCTAACTGTACCAAAATAGACTAGCTTTAATAAAATCTT	Majority
	1410 1420 1430 1440 1450	
1302	AGATACGGTCAAGCTAACTGTACCAAAATAGACTAGCTTTAATAAAATCTT	2603_a12.seq
1401	AGATACGGTCAAGCTAACTGTACCAAAATAGACTAGCTTTAATAAAATCTT	nem316_a12.seq
	TTGCACTCTCTCTATTTTTTCCAGAAAATAGCGAAACTTCTTAAAAATAAA	Majority
	1460 1470 1480 1490 1500	
1352	TTGCACTCTCTCTATTTTTTCCAGAAAATAGCGAAACTTCTTAAAAATAAA	2603_a12.seq
1451	TTGCACTCTCTCTATTTTTTCCAGAAAATAGCGAAACTTCTTAAAAATAAA	nem316_a12.seq
	GCTAGAGCAACCATATTCATCGGTAAACCGATAAAGGTTTCTGGACCACG	Majority
	1510 1520 1530 1540 1550	
1402	GCTAGAGCAACCATATTCATCGGTAAACCGATAAAGGTTTCTGGACCACG	2603_a12.seq
1501	GCTAGAGCAACCATATTCATCGGTAAACCGATAAAGGTTTCTGGACCACG	nem316_a12.seq
	ATTAGCAAGTATAAAGTTTAAAGTGTCTTAATAAGAGTACACCATAAC	Majority
	1560 1570 1580 1590 1600	
1452	ATTAGCAAGTATAAAGTTTAAAGTGTCTTAATAAGAGTACACCATAAC	2603_a12.seq
1551	ATTAGCAAGTATAAAGTTTAAAGTGTCTTAATAAGAGTACACCATAAC	nem316_a12.seq
	TTGATTTTCAAATCAAATAAAATAAAGCAACTAACATCGGAAGGATTGAA	Majority
	1610 1620 1630 1640 1650	
1502	TTGATTTTCAAATCAAATAAAATAAAGCAACTAACATCGGAAGGATTGAA	2603_a12.seq
1601	TTGATTTTCAAATCAAATAAAATAAAGCAACTAACATCGGAAGGATTGAA	nem316_a12.seq
	AAATCAACCTTTAAAAAATTCTGCTCCTGGTATTAATGGAAATGAAACCAT	Majority
	1660 1670 1680 1690 1700	
1552	AAATCAACCTTTAAAAAATTCTGCTCCTGGTATTAATGGAAATGAAACCAT	2603_a12.seq
1651	AAATCAACCTTTAAAAAATTCTGCTCCTGGTATTAATGGAAATGAAACCAT	nem316_a12.seq
	CATCAATACAAAAGATAAGGCAGAAAGAATGGCGATTGTCAACCATTTTAC	Majority
	1710 1720 1730 1740 1750	
1602	CATCAATACAAAAGATAAGGCAGAAAGAATGGCGATTGTCAACCATTTTAC	2603_a12.seq
1701	CATCAATACAAAAGATAAGGCAGAAAGAATGGCGATTGTCAACCATTTTAC	nem316_a12.seq
	GTCGATTTTGTATATAAAAAAATTCCTCCAATTTAAATAAATTGAAAGAAAGC	Majority
	1760 1770 1780 1790 1800	
1652	GTCGATTTTGTATATAAAAAAATTCCTCCAATTTAAATAAATTGAAAGAAAGC	2603_a12.seq
1751	GTCGATTTTGTATATAAAAAAATTCCTCCAATTTAAATAAATTGAAAGAAAGC	nem316_a12.seq
	TCCAAAGGTAAGCGTATGTACGGCGAAAAAACCTTGTCTTCTCCCATCC	Majority
	1810 1820 1830 1840 1850	
1702	TCCAAAGGTAAGCGTATGTACGGCGAAAAAACCTTGTCTTCTCCCATCC	2603_a12.seq
1801	TCCAAAGGTAAGCGTATGTACGGCGAAAAAACCTTGTCTTCTCCCATCC	nem316_a12.seq
	AGACTTTTACTGTCGGTTGTGGAATCTCACCACATCAGCTTTTCGCTCGCGG	Majority
	1860 1870 1880 1890 1900	
1751	AGACTTTTACTGTCGGTTGTGGAATCTCACCACATCAGCTTTTCGCTCGCGG	2603_a12.seq
1851	AGACTTTTACTGTCGGTTGTGGAATCTCACCACATCAGCTTTTCGCTCGCGG	nem316_a12.seq
	ACTGATGCTTTCACAACTGACAAAATAAGTTGGAAGCGATTACCGCCGGTCC	Majority
	1910 1920 1930 1940 1950	
1801	ACTGATGCTTTCACAACTGACAAAATAAGTTGGAAGCGATTACCGCCGGTCC	2603_a12.seq
1901	ACTGATGCTTTCACAACTGACAAAATAAGTTGGAAGCGATTACCGCCGGTCC	nem316_a12.seq

FIGURE 20B

91/487

		GGAATTACACCCCTGCCCTGAAGACACCTATAGCATAACAAAAAACTTG					Majority
		1960	1970	1980	1990	2000	
1851		GGAATTACACCCCTGCCCTGAAGACACCTATAGCATAACAAAAAACTTG					2603_a12.seq
1951		GGAATTACACCCCTGCCCTGAAGACACCTATAGCATAACAAAAAACTTG					nem316_a12.seq
		CAATTGCAAGTTTTTTTAATTACTAATTAGTAGTAGTGATTAAAAATCATA					Majority
		2010	2020	2030	2040	2050	
1901		CAATTGCAAGTTTTTTTAATTACTAATTAGTAGTAGTGATTAAAAATCATA					2603_a12.seq
2001		CAATTGCAAGTTTTTTTAATTACTAATTAGTAGTAGTGATTAAAAATCATA					nem316_a12.seq
		TTAATACCAAATTACTATGCTGTATCGTTTCTTTTCAGATTGCTATTTTT					Majority
		2060	2070	2080	2090	2100	
1951		TTAATACCAAATTACTATGCTGTATCGTTTCTTTTCAGATTGCTATTTTT					2603_a12.seq
2051		TTAATACCAAATTACTATGCTGTATCGTTTCTTTTCAGATTGCTATTTTT					nem316_a12.seq
		AGTTTTTCTTAAAAAGATAAAACAAAATTCCCAAAATAATACAACCAAGAA					Majority
		2110	2120	2130	2140	2150	
2001		AGTTTTTCTTAAAAAGATAAAACAAAATTCCCAAAATAATACAACCAAGAA					2603_a12.seq
2101		AGTTTTTCTTAAAAAGATAAAACAAAATTCCCAAAATAATACAACCAAGAA					nem316_a12.seq
		TTGTCAGTCCCTCCACCAATAATCATTCCTGTTTTAGGAAGAAATGATTGT					Majority
		2160	2170	2180	2190	2200	
2051		TTGTCAGTCCCTCCACCAATAATCATTCCTGTTTTAGGAAGAAATGATTGT					2603_a12.seq
2151		TTGTCAGTCCCTCCACCAATAATCATTCCTGTTTTAGGAAGAAATGATTGT					nem316_a12.seq
		GGAAAAAGCGGTTGTCATCGTTTAGGATTTGTTGCTGGAAGAGTTTCTTT					Majority
		2210	2220	2230	2240	2250	
2101		GGAAAAAGCGGTTGTCATCGTTTAGGATTTGTTGCTGGAAGAGTTTCTTT					2603_a12.seq
2201		GGAAAAAGCGGTTGTCATCGTTTAGGATTTGTTGCTGGAAGAGTTTCTTT					nem316_a12.seq
		TTCGTTTTCTACCTCTACTTCTCTGTGTTTTATTAGCAACTACAGCAACTA					Majority
		2260	2270	2280	2290	2300	
2151		TTCGTTTTCTACCTCTACTTCTCTGTGTTTTATTAGCAACTACAGCAACTA					2603_a12.seq
2251		TTCGTTTTCTACCTCTACTTCTCTGTGTTTTATTAGCAACTACAGCAACTA					nem316_a12.seq
		CAGCATCCTTCATAGATATACGGTAACCAGTTAGTGCTTTTGCTTCTCGA					Majority
		2310	2320	2330	2340	2350	
2201		CAGCATCCTTCATAGATATACGGTAACCAGTTAGTGCTTTTGCTTCTCGA					2603_a12.seq
2301		CAGCATCCTTCATAGATATACGGTAACCAGTTAGTGCTTTTGCTTCTCGA					nem316_a12.seq
		AAAAATATACTTACCAGGTAATAAACCTTCAACCTCAATTTCTCCCTTATC					Majority
		2360	2370	2380	2390	2400	
2251		AAAAATATACTTACCAGGTAATAAACCTTCAACCTCAATTTCTCCCTTATC					2603_a12.seq
2351		AAAAATATACTTACCAGGTAATAAACCTTCAACCTCAATTTCTCCCTTATC					nem316_a12.seq
		ATCAGTTACTAATGAAGTAATCCCATCTTGATCGGTCGTAATTCGTCCTAT					Majority
		2410	2420	2430	2440	2450	
2301		ATCAGTTACTAATGAAGTAATCCCATCTTGATCGGTCGTAATTCGTCCTAT					2603_a12.seq
2401		ATCAGTTACTAATGAAGTAATCCCATCTTGATCGGTCGTAATTCGTCCTAT					nem316_a12.seq
		TTTTAAAGCGAACTGGCTGATTCTGTTATCGTATAATAACAATATTACT					Majority
		2460	2470	2480	2490	2500	
2351		TTTTAAAGCGAACTGGCTGATTCTGTTATCGTATAATAACAATATTACT					2603_a12.seq
2451		TTTTAAAGCGAACTGGCTGATTCTGTTATCGTATAATAACAATATTACT					nem316_a12.seq
		CCTGATAGCCTTTTCTTTATCTTTCTTTCTTTTGTATATTTAATAAGTTT					Majority
		2510	2520	2530	2540	2550	
2401		CCTGATAGCCTTTTCTTTATCTTTCTTTCTTTTGTATATTTAATAAGTTT					2603_a12.seq
2501		CCTGATAGCCTTTTCTTTATCTTTCTTTCTTTTGTATATTTAATAAGTTT					nem316_a12.seq
		TAATCGGCCTGTTTCAACTTTTTCGCTTAGGATTTATCTGTAATTGATTTG					Majority
		2560	2570	2580	2590	2600	
2451		TAATCGGCCTGTTTCAACTTTTTCGCTTAGGATTTATCTGTAATTGATTTG					2603_a12.seq
2551		TAATCGGCCTGTTTCAACTTTTTCGCTTAGGATTTATCTGTAATTGATTTG					nem316_a12.seq

FIGURE 20C

		ATAACTTATCATCTGGTATTTCAATATAAAAAAGGTACTATTGTTGAAACG Majority				
		2610	2620	2630	2640	2650
2501	2601	ATAACTTATCATCTGGTATTTCAATATAAAAAAGGTACTATTGTTGAAACG 2603_al2.seq				
		ATAACTTATCATCTGGTATTTCAATATAAAAAAGGTACTATTGTTGAAACG nem316_al2.seq				
		CTTTGATCAGCTTTATAAGCACGACCAAAGTACGAACCATTTGGGAGTGC Majority				
		2660	2670	2680	2690	2700
2551	2651	CTTTGATCAGCTTTATAAGCACGACCAAAGTACGAACCATTTGGGAGTGC 2603_al2.seq				
		CTTTGATCAGCTTTATAAGCACGACCAAAGTACGAACCATTTGGGAGTGC nem316_al2.seq				
		TATCTTTGCTCTGACCATTAGTATCAGTAGGAGAAGTCAAGATACTCTTAT Majority				
		2710	2720	2730	2740	2750
2601	2701	TATCTTTGCTCTGACCATTAGTATCAGTAGGAGAAGTCAAGATACTCTTAT 2603_al2.seq				
		TATCTTTGCTCTGACCATTAGTATCAGTAGGAGAAGTCAAGATACTCTTAT nem316_al2.seq				
		ACTTCTGCTTCAATTGCTATCTGTCAATTTGGCTCAATAAATCAACTTTT Majority				
		2760	2770	2780	2790	2800
2651	2751	ACTTCTGCTTCAATTGCTATCTGTCAATTTGGCTCAATAAATCAACTTTT 2603_al2.seq				
		ACTTCTGCTTCAATTGCTATCTGTCAATTTGGCTCAATAAATCAACTTTT nem316_al2.seq				
		AAGTTGTCAGTCACAGTCCATAAACGATAAGAAATCCCCTCCTCTGTAGT Majority				
		2810	2820	2830	2840	2850
2701	2801	AAGTTGTCAGTCACAGTCCATAAACGATAAGAAATCCCCTCCTCTGTAGT 2603_al2.seq				
		AAGTTGTCAGTCACAGTCCATAAACGATAAGAAATCCCCTCCTCTGTAGT nem316_al2.seq				
		ATTTGGCTGAAGTCCCTATCTGTGTGATTGTTAGTTGATTAGGGGTATCAG Majority				
		2860	2870	2880	2890	2900
2751	2851	ATTTGGCTGAAGTCCCTATCTGTGTGATTGTTAGTTGATTAGGGGTATCAG 2603_al2.seq				
		ATTTGGCTGAAGTCCCTATCTGTGTGATTGTTAGTTGATTAGGGGTATCAG nem316_al2.seq				
		CATTTACACTGGCTACCGAAAAAAACGCTAATTGTACCAATCCTAAAAAG Majority				
		2910	2920	2930	2940	2950
2801	2901	CATTTACACTGGCTACCGAAAAAAACGCTAATTGTACCAATCCTAAAAAG 2603_al2.seq				
		CATTTACACTGGCTACCGAAAAAAACGCTAATTGTACCAATCCTAAAAAG nem316_al2.seq				
		CAACATAGTAGAAGTCCTAAACCTTTTCTAATCTTTTTTCATTTTTTGATT Majority				
		2960	2970	2980	2990	3000
2851	2951	CAACATAGTAGAAGTCCTAAACCTTTTCTAATCTTTTTTCATTTTTTGATT 2603_al2.seq				
		CAACATAGTAGAAGTCCTAAACCTTTTCTAATCTTTTTTCATTTTTTGATT nem316_al2.seq				
		CCCTTTCTTTTTCTCTCTTTTAAATTTTTCGTTTTTAAATATAATAGTAAAGC Majority				
		3010	3020	3030	3040	3050
2901	3001	CCCTTTCTTTTTCTCTCTTTTAAATTTTTCGTTTTTAAATATAATAGTAAAGC 2603_al2.seq				
		CCCTTTCTTTTTCTCTCTTTTAAATTTTTCGTTTTTAAATATAATAGTAAAGC nem316_al2.seq				
		GACTAATATAAGAATAAAGTATGATAAGAGGAAATAAAGTTTATAGT Majority				
		3060	3070	3080	3090	3100
2951	3051	GACTAATATAAGAATAAAGTATGATAAGAGGAAATAAAGTTTATAGT 2603_al2.seq				
		GACTAATATAAGAATAAAGTATGATAAGAGGAAATAAAGTTTATAGT nem316_al2.seq				
		GTGTTTGCAATTGTTTCAATTAATAAGTTCTTTTCTTTTAAACAGGAGGTACA Majority				
		3110	3120	3130	3140	3150
3001	3101	GTGTTTGCAATTGTTTCAATTAATAAGTTCTTTTCTTTTAAACAGGAGGTACA 2603_al2.seq				
		GTGTTTGCAATTGTTTCAATTAATAAGTTCTTTTCTTTTAAACAGGAGGTACA nem316_al2.seq				
		TACTTGATTCCGATGCCCTCTAAGTAAACGATGTGAATTAATCGAATA Majority				
		3160	3170	3180	3190	3200
3051	3151	TACTTGATTCCGATGCCCTCTAAGTAAACGATGTGAATTAATCGAATA 2603_al2.seq				
		TACTTGATTCCGATGCCCTCTAAGTAAACGATGTGAATTAATCGAATA nem316_al2.seq				
		AGGTGTACATGTTAGCAAAGTCGCATAATCCTTACCTTTAACAACCAATA Majority				
		3210	3220	3230	3240	3250
3101	3201	AGGTGTACATGTTAGCAAAGTCGCATAATCCTTACCTTTAACAACCAATA 2603_al2.seq				
		AGGTGTACATGTTAGCAAAGTCGCATAATCCTTACCTTTAACAACCAATA nem316_al2.seq				

FIGURE 20D

93/487

```

      3260      3270      3280      3290      3300
ATTATGAAATAATTATCTGGCTTTTACAACTTATTGATCAACCTTATAG Majority
3151 ATTTAGAAAAATTATCTGGCTTTTACAACTTATTGATCAACCTTATAG 2603_al2.seq
3251 ATTTAGAAAAATTATCTGGCTTTTACAACTTATTGATCAACCTTATAG nem316_al2.seq

      3310      3320      3330      3340      3350
GCTAAAACTTCTTTGATATTATGAATATAAAAAATTTTCTTTTTTAAAG Majority
3201 GCTAAAACTTCTTTGATATTATGAATATAAAAAATTTTCTTTTTTAAAG 2603_al2.seq
3301 GCTAAAACTTCTTTGATATTATGAATATAAAAAATTTTCTTTTTTAAAG nem316_al2.seq

      3360      3370      3380      3390      3400
TTTATCTAAATCTGTAAATAACTTAGCTTTAGGTAAGCCCGGATGAGCTG Majority
3251 TTTATCTAAATCTGTAAATAACTTAGCTTTAGGTAAGCCCGGATGAGCTG 2603_al2.seq
3351 TTTATCTAAATCTGTAAATAACTTAGCTTTAGGTAAGCCCGGATGAGCTG nem316_al2.seq

      3410      3420      3430      3440      3450
TGATAACAGTATGTGAACCTTTTCCACCAATTGGCAAGGAGGTTCTTCA Majority
3301 TGATAACAGTATGTGAACCTTTTCCACCAATTGGCAAGGAGGTTCTTCA 2603_al2.seq
3401 TGATAACAGTATGTGAACCTTTTCCACCAATTGGCAAGGAGGTTCTTCA nem316_al2.seq

      3460      3470      3480      3490      3500
AGGTGTCCTGCTCCTTTTTTCAAGAACACTACTGCTAGTCCCGCATAGAT Majority
3351 AGGTGTCCTGCTCCTTTTTTCAAGAACACTACTGCTAGTCCCGCATAGAT 2603_al2.seq
3451 AGGTGTCCTGCTCCTTTTTTCAAGAACACTACTGCTAGTCCCGCATAGAT nem316_al2.seq

      3510      3520      3530      3540      3550
AGGTAATTTTTGCTTGATAGACGGTATATCAATATATCCAATCATTTCAG Majority
3401 AGGTAATTTTTGCTTGATAGACGGTATATCAATATATCCAATCATTTCAG 2603_al2.seq
3501 AGGTAATTTTTGCTTGATAGACGGTATATCAATATATCCAATCATTTCAG nem316_al2.seq

      3560      3570      3580      3590      3600
TAATCTCAAGCATGTGGGCGTATTCAGCAATACCTTTTTTCTTTTTTCA Majority
3451 TAATCTCAAGCATGTGGGCGTATTCAGCAATACCTTTTTTCTTTTTTCA 2603_al2.seq
3551 TAATCTCAAGCATGTGGGCGTATTCAGCAATACCTTTTTTCTTTTTTCA nem316_al2.seq

      3610      3620      3630      3640      3650
GTATAGGGATCTGATAGGCGGCTTGGGTCCAGTGTCTATTATAAGCTTT Majority
3501 GTATAGGGATCTGATAGGCGGCTTGGGTCCAGTGTCTATTATAAGCTTT 2603_al2.seq
3601 GTATAGGGATCTGATAGGCGGCTTGGGTCCAGTGTCTATTATAAGCTTT nem316_al2.seq

      3660      3670      3680      3690      3700
TGCTAACTCAAATCGTCTATTAATCTCTTTAGTATTTAATTTTTGGGTTT Majority
3551 TGCTAACTCAAATCGTCTATTAATCTCTTTAGTATTTAATTTTTGGGTTT 2603_al2.seq
3651 TGCTAACTCAAATCGTCTATTAATCTCTTTAGTATTTAATTTTTGGGTTT nem316_al2.seq

      3710      3720      3730      3740      3750
GATTATCAAAGTTAGTTACTTGATTATTAGCTTTAATATTATAGTACCAA Majority
3601 GATTATCAAAGTTAGTTACTTGATTATTAGCTTTAATATTATAGTACCAA 2603_al2.seq
3701 GATTATCAAAGTTAGTTACTTGATTATTAGCTTTAATATTATAGTACCAA nem316_al2.seq

      3760      3770      3780      3790      3800
TTTGAAATAAAAGGATATGAGGTTATCAAAAAGACCAACTAAGAACAAATAG Majority
3651 TTTGAAATAAAAGGATATGAGGTTATCAAAAAGACCAACTAAGAACAAATAG 2603_al2.seq
3751 TTTGAAATAAAAGGATATGAGGTTATCAAAAAGACCAACTAAGAACAAATAG nem316_al2.seq

      3810      3820      3830      3840      3850
TATCAGGCCTACATTTCATCCATCGATTTTAAAAAGACCGATTTCCTTAAGCT Majority
3701 TATCAGGCCTACATTTCATCCATCGATTTTAAAAAGACCGATTTCCTTAAGCT 2603_al2.seq
3801 TATCAGGCCTACATTTCATCCATCGATTTTAAAAAGACCGATTTCCTTAAGCT nem316_al2.seq

      3860      3870      3880      3890      3900
TTTTCTGAAATTTTCTCCCATTTATGATTCAATTCCTTTTCTAACACTTG Majority
3751 TTTTCTGAAATTTTCTCCCATTTATGATTCAATTCCTTTTCTAACACTTG 2603_al2.seq
3851 TTTTCTGAAATTTTCTCCCATTTATGATTCAATTCCTTTTCTAACACTTG nem316_al2.seq

```

FIGURE 20E

94/487

	CTAAA	3910	3920	3930	3940	3950	Majority
3801	CTAAACGATTTTTTTTGACGTTGACGTTTTTATTAACCAAAGTAACCAAGCA						
3901	CTAAACGATTTTTTTTGACGTTGACGTTTTTATTAACCAAAGTAACCAAGCA						2603_al2.seq nem316_al2.seq
	ATAATAACTAAAGATATATAGAAATAGATATCTATAAATCGTGTTTAAATG	3960	3970	3980	3990	4000	Majority
3851	ATAATAACTAAAGATATATAGAAATAGATATCTATAAATCGTGTTTAAATG						
3951	ATAATAACTAAAGATATATAGAAATAGATATCTATAAATCGTGTTTAAATG						2603_al2.seq nem316_al2.seq
	ACCTTCTTTTTATTAATTTTTTCATCAATAGGACCTTTATAAGGGATACGAT	4010	4020	4030	4040	4050	Majority
3901	ACCTTCTTTTTATTAATTTTTTCATCAATAGGACCTTTATAAGGGATACGAT						
4001	ACCTTCTTTTTATTAATTTTTTCATCAATAGGACCTTTATAAGGGATACGAT						2603_al2.seq nem316_al2.seq
	GTCCCCCTTACTAAAAGTCTGTGTGTATTGATCATAATCGGGGTGCAAGTT	4060	4070	4080	4090	4100	Majority
3951	GTCCCCCTTACTAAAAGTCTGTGTGTATTGATCATAATCGGGGTGCAAGTT						
4051	GTCCCCCTTACTAAAAGTCTGTGTGTATTGATCATAATCGGGGTGCAAGTT						2603_al2.seq nem316_al2.seq
	AATAAGGTTGCCATAATCATGTCCAGGAACAACCAACAAATCTGAAAAGTT	4110	4120	4130	4140	4150	Majority
4001	AATAAGGTTGCCATAATCATGTCCAGGAACAACCAACAAATCTGAAAAGTT						
4101	AATAAGGTTGCCATAATCATGTCCAGGAACAACCAACAAATCTGAAAAGTT						2603_al2.seq nem316_al2.seq
	ATCGGGTGTAACGACTTTTTATCTGATCTACTTGATATGCTATCGTTTCTT	4160	4170	4180	4190	4200	Majority
4051	ATCGGGTGTAACGACTTTTTATCTGATCTACTTGATATGCTATCGTTTCTT						
4151	ATCGGGTGTAACGACTTTTTATCTGATCTACTTGATATGCTATCGTTTCTT						2603_al2.seq nem316_al2.seq
	TTATGTTTTTGAATATAAAACTTATCTCCTTTTTTTTAACTTTTTAAGGTTA	4210	4220	4230	4240	4250	Majority
4101	TTATGTTTTTGAATATAAAACTTATCTCCTTTTTTTTAACTTTTTAAGGTTA						
4201	TTATGTTTTTGAATATAAAACTTATCTCCTTTTTTTTAACTTTTTAAGGTTA						2603_al2.seq nem316_al2.seq
	GAAAAGAGTTCTTTTATCTGGAATTCCTGAGTGCCTGTTATAACGGTATG	4260	4270	4280	4290	4300	Majority
4151	GAAAAGAGTTCTTTTATCTGGAATTCCTGAGTGCCTGTTATAACGGTATG						
4251	GAAAAGAGTTCTTTTATCTGGAATTCCTGAGTGCCTGTTATAACGGTATG						2603_al2.seq nem316_al2.seq
	TGTGCTATTTTCTTCCAATTGGAAGAGAGGTACCTTCTAAATGCCCTGCTC	4310	4320	4330	4340	4350	Majority
4201	TGTGCTATTTTCTTCCAATTGGAAGAGAGGTACCTTCTAAATGCCCTGCTC						
4301	TGTGCTATTTTCTTCCAATTGGAAGAGAGGTACCTTCTAAATGCCCTGCTC						2603_al2.seq nem316_al2.seq
	CTTTAGATAGAACTTCTTGACTTGAACCTGCAAAATATAGGGAGTTTTTGA	4360	4370	4380	4390	4400	Majority
4251	CTTTAGATAGAACTTCTTGACTTGAACCTGCAAAATATAGGGAGTTTTTGA						
4351	CTTTAGATAGAACTTCTTGACTTGAACCTGCAAAATATAGGGAGTTTTTGA						2603_al2.seq nem316_al2.seq
	CCTATCTTAGGAACTGAAATTGTTCCGATTTTTTTCACTTACCTCTAACAT	4410	4420	4430	4440	4450	Majority
4391	CCTATCTTAGGAACTGAAATTGTTCCGATTTTTTTCACTTACCTCTAACAT						
4401	CCTATCTTAGGAACTGAAATTGTTCCGATTTTTTTCACTTACCTCTAACAT						2603_al2.seq nem316_al2.seq
	ACGGGGCGTACTCTGCTACCCCTTTTGAATTTCGTTTTTCTCATAAGGAT	4460	4470	4480	4490	4500	Majority
4351	ACGGGGCGTACTCTGCTACCCCTTTTGAATTTCGTTTTTCTCATAAGGAT						
4451	ACGGGGCGTACTCTGCTACCCCTTTTGAATTTCGTTTTTCTCATAAGGAT						2603_al2.seq nem316_al2.seq
	CTTCAAGATGGACATTATTTAAAGAATCATTATAAGCTTGTGCTAGAGTC	4510	4520	4530	4540	4550	Majority
4401	CTTCAAGATGGACATTATTTAAAGAATCATTATAAGCTTGTGCTAGAGTC						
4501	CTTCAAGATGGACATTATTTAAAGAATCATTATAAGCTTGTGCTAGAGTC						2603_al2.seq nem316_al2.seq

FIGURE 20F

95/487

	ATACGTCGCAATPGATTTCTCTCACTTAAGTTTTTTTAGCAGCTCTCTCTCAAA	Majority
	4560 4570 4580 4590 4600	
4451	ATACGTCGATTCGATTTCTTTCTGACTAAGTTTTTTTAGCAGCTCTCTCTCAAA	2603_a12.seq
4551	ATACGTCGATTCGATTTCTTTCTGACTAAGTTTTTTTAGCAGCTCTCTCTCAAA	nem316_a12.seq
	ATCCTGTGTTTGATTATTAGATTCTATCGTATAGTAAAAACGTGATACCA	Majority
	4610 4620 4630 4640 4650	
4501	ATCCTGTGTTTGATTATTAGATTCTATCGTATAGTAAAAACGTGATACCA	2603_a12.seq
4601	ATCCTGTGTTTGATTATTAGATTCTATCGTATAGTAAAAACGTGATACCA	nem316_a12.seq
	CTGGATACAATAAAAATAGATAGACCTATTAGAAAAAGAATGATAAAAAGGA	Majority
	4660 4670 4680 4690 4700	
4551	CTGGATACAATAAAAATAGATAGACCTATTAGAAAAAGAATGATAAAAAGGA	2603_a12.seq
4651	CTGGATACAATAAAAATAGATAGACCTATTAGAAAAAGAATGATAAAAAGGA	nem316_a12.seq
	AGATTTGACTTCTTCTTTTTTTTTTTGTTTTTTTGTTCATTTTTTTACTCTT	Majority
	4710 4720 4730 4740 4750	
4601	AGATTTGACTTCTTCTTTTTTTTTTTGTTTTTTTGTTCATTTTTTTACTCTT	2603_a12.seq
4701	AGATTTGACTTCTTCTTTTTTTTAA...GTTTTTTTGTTCATTTTTTTACTCTT	nem316_a12.seq
	CACGTCATCTCCTAGATAATGGCTCTTGCCTTATGATCTAAGAGTACTTCT	Majority
	4760 4770 4780 4790 4800	
4651	CACGTCATCTCCTAGATAATGGCTCTTGCCTTATGATCTAAGAGTACTTCT	2603_a12.seq
4748	CACGTCATCTCCTAATAAATGGCTCTTGCCTTATGATCTAAGAGTACTTCT	nem316_a12.seq
	ACTGAAATACCCCTTAGATCATAAGCACAGCTTTAACTGTGCTTATACATC	Majority
	4810 4820 4830 4840 4850	
4701	ACTGAAATACCCCTTAGATCATAAGCACAGCTTTAACTGTGCTTATACATC	2603_a12.seq
4798	ACTGAAATACCCCTTAGATCATAAGCACAGCTTTAACTGTGCTTATACATC	nem316_a12.seq
	ATCAAAGACTAGCCCTTAAGCTTCCTTTGATTGGCGTTTTTTTCATGATAAC	Majority
	4860 4870 4880 4890 4900	
4751	ATCAAAGACTAGCCCTTAAGCTTCCTTTGATTGGCGTTTTTTTCATGATAAC	2603_a12.seq
4848	ATCAAAGACTAGCCCTTAAGCTTCCTTTGATTGGCGTTTTTTTCATGATAAC	nem316_a12.seq
	TACTGCTCCAAGCATAAATGCTTAAACCAATAATTGTGAAAAGAATTGTAC	Majority
	4910 4920 4930 4940 4950	
4801	TACTGCTCCAAGCATAAATGCTTAAACCAATAATTGTGAAAAGAATTGTAC	2603_a12.seq
4898	TACTGCTCCAAGCATAAATGCTTAAACCAATAATTGTGAAAAGAATTGTAC	nem316_a12.seq
	CAATACCACCTGTTTTGTGGGATTGTTACTTTTTTGTTTTGTACTTGTTTG	Majority
	4960 4970 4980 4990 5000	
4851	CAATACCACCTGTTTTGTGGGATTGTTACTTTTTTGTTTTGTACTTGTTTG	2603_a12.seq
4948	CAATACCACCTGTTTTGTGGGATTGTTACTTTTTTGTTTTGTACTTGTTTG	nem316_a12.seq
	GCATCTTTTTTTTACAGGTTTTTTTGTATCTGCGTTGTCAGTTTTAGCCCC	Majority
	5010 5020 5030 5040 5050	
4901	GCATCTTTTTTTTACAGGTTTTTTTGTATCTGCGTTGTCAGTTTTAGCCCC	2603_a12.seq
4998	GCATCTTTTTTTTACAGGTTTTTTTGTATCTGCGTTGTCAGTTTTAGCCCC	nem316_a12.seq
	TTTTCTGTATGATGTTTGAATTACTTCAAAGTTTATATTAACCTGCCAATT	Majority
	5060 5070 5080 5090 5100	
4945	TTTTCTGTATGATGTTTGAATTACTTCAAAGTTTATATTAACCTGCCAATT	2603_a12.seq
5048	TTTTCTGTATGATGTTTGAATTACTTCAAAGTTTATATTAACCTGCCAATT	nem316_a12.seq
	TCCGATATCCTGCTGCTGCTTGTGTTTTCTTCCAGGTTGTAAGTGCCTTTT	Majority
	5110 5120 5130 5140 5150	
4995	TCCGATATCCTGCTGCTGCTTGTGTTTTCTTCCAGGTTGTAAGTGCCTTTT	2603_a12.seq
5098	TCCGATATCCTGCTGCTGCTTGTGTTTTCTTCCAGGTTGTAAGTGCCTTTT	nem316_a12.seq
	TCCAGACCTGTAAATTTCAAATTGACCTTGGTTCGTTTCAGGTGTAATTTAAT	Majority
	5160 5170 5180 5190 5200	
5045	TCCAGACCTGTAAATTTCAAATTGACCTTGGTTCGTTTCAGGTGTAATTTAAT	2603_a12.seq
5148	TCCAGACCTGTAAATTTCAAATTGACCTTGGTTCGTTTCAGGTGTAATTTAAT	nem316_a12.seq

FIGURE 20G

FIGURE 20H



98/487

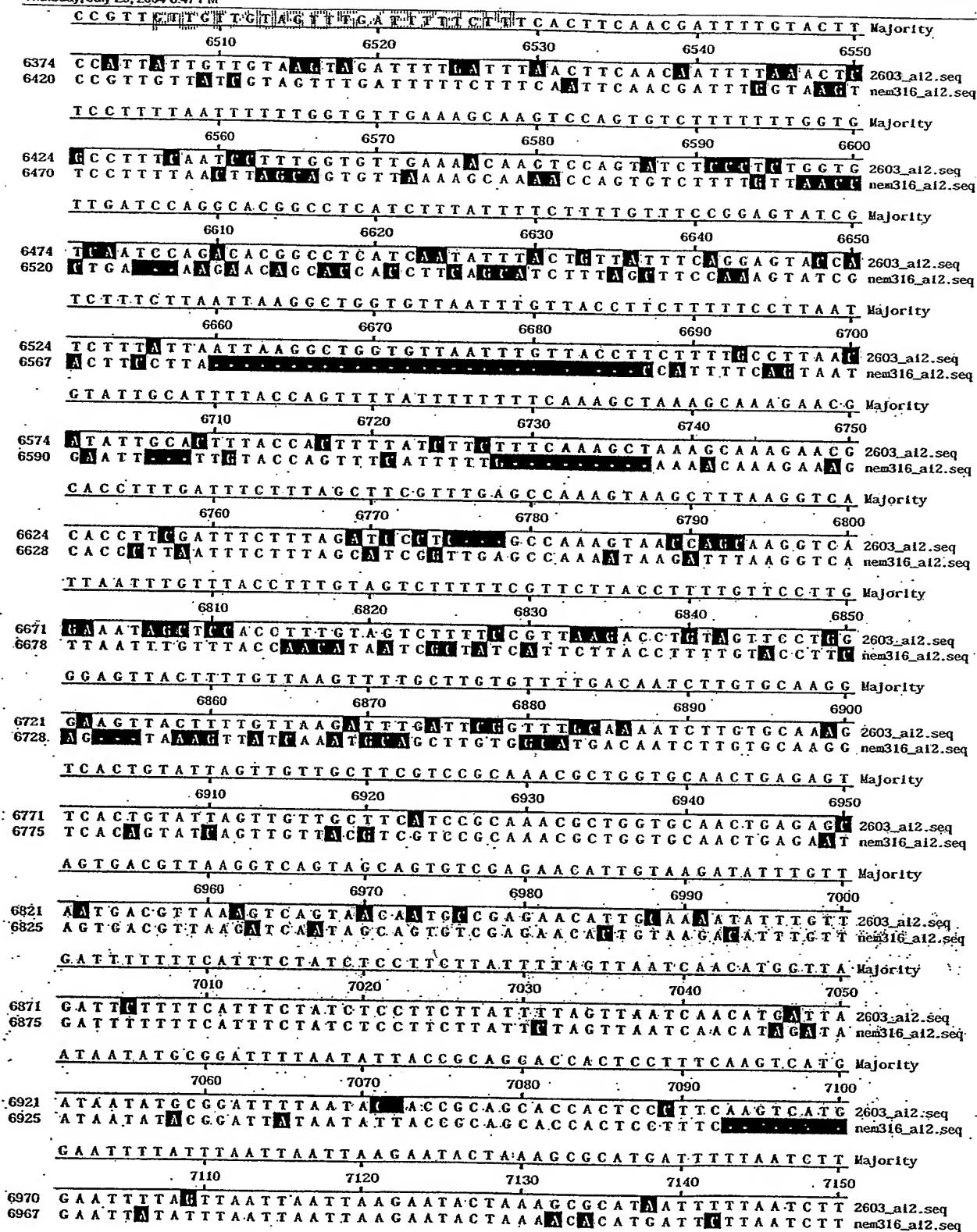


FIGURE 20J

99/487

	TTTTTCTGGCAATATATGAGTAACTTCTTATATCTTTTCCAAATATAAAATT	Majority
	7160 7170 7180 7190 7200	
7020	TTTTCTGGCAATATATCACTAGATTCTTATACTTTTCCAAATATAAAATT	2603_al2.seq
7017	TTTTCTCTAGCATATATCACTAGATTCTTATACTTTTCCAAATATAAAATT	nem316_al2.seq
	GCACCTGCAATAGACATCATAGCTCCACCTATTAAAAATGAAAGATAGAAT	Majority
	7210 7220 7230 7240 7250	
7070	CCACCTGCAATAGACATCATAGCTCCACCTATTAAAAATGAAAGATAGAAT	2603_al2.seq
7067	CCACCTGCAATAGACATCATAGCTCCACCTATTAAAAATGAAAGATAGAAT	nem316_al2.seq
	TCCTTTCCCACCTGTCATCGGAATAATTCTTTTGGTGGAAATATGCGTGT	Majority
	7260 7270 7280 7290 7300	
7120	TGCTTTCCCACCTGTCATCGGAATAATTCTTTTGGTGGAAATATGCGTGT	2603_al2.seq
7117	TCCTTTCCCACCTGTCATCGGAATAATTCTTTTGGTGGAAATATGCGTGT	nem316_al2.seq
	TGGTAATTAATGCTTGTACCTTCCCTCATGATATTCAGAAATCTGTTTA	Majority
	7310 7320 7330 7340 7350	
7170	TGGTAATTAATGCTTGTACCTTCCCTCATGATATTCAGAAATCTGTTTA	2603_al2.seq
7167	TGGTAATTAATGCTTGTACCTTCCCTCATGATATTCAGAAATCTGTTTA	nem316_al2.seq
	TAAACAGCTATTATATTTTTTATCGATCCTTTAACCACCTTCAAAAAGTTAA	Majority
	7360 7370 7380 7390 7400	
7220	TAAACAGCTATTATATTTTTTATCGATCCTTTAACCACCTTCAAAAAGTTAA	2603_al2.seq
7217	TAAACAGCTATTATATTTTTTATCGATCCTTTAACAACCTTCAAAAAGTTAA	nem316_al2.seq
	AATTGGTTTATTAGTAATTTTTTGATAATCCTTCGGCGAAACTGCTTCTA	Majority
	7410 7420 7430 7440 7450	
7270	AATTGGTTTATTAGTAATTTTTTGATAATCCTTCGGCGAAACTGCTTCTA	2603_al2.seq
7267	AATTGGTTTATTAGTAATTTTTTGATAATCCTTCGGCGAAACTGCTTCTA	nem316_al2.seq
	TAACTGATATTTGCCATCTTTCAAATCTTTGTAAGAAATTTTGCCCGTTT	Majority
	7460 7470 7480 7490 7500	
7320	TAACTGATATTTGCCATCTTTCAAATCTTTGTAAGAAATTTTGCCCGTTT	2603_al2.seq
7317	TAACTGATATTTGCCATCTTTCAAATCTTTGTAAGAAATTTTGCCCGTTT	nem316_al2.seq
	TCTCCCGTCACTACTTTTGAATTATTATTTTTTATTGCTAAATAAAAGTTT	Majority
	7510 7520 7530 7540 7550	
7370	TCTCCCGTCACTACTTTTGAATTATTATTTTTTATTGCTAAATAAAAGTTT	2603_al2.seq
7367	TCTCCCGTCACTACTTTTGAATTATTATTTTTTATTGCTAAATAAAAGTTT	nem316_al2.seq
	ATAATCTTCATTAAATTCCTTGAAGTTCAAACGCTAGCTCCTTTGAGAAGCA	Majority
	7560 7570 7580 7590 7600	
7420	ATAATCTTCATTAAATTCCTTGAAGTTCAAACGCTAGCTCCTTTGAGAAGCA	2603_al2.seq
7417	ATAATCTTCATTAAATTCCTTGAAGTTCAAACGCTAGCTCCTTTGAGAAGCA	nem316_al2.seq
	ACTTATTATTATCTTTATCAACTTTTGTAAATTCAAATTTACCTAACTTC	Majority
	7610 7620 7630 7640 7650	
7470	ACTTATTATTATCTTTATCAACTTTTGTAAATTCAAATTTACCTAACTTC	2603_al2.seq
7467	ACTTATTATTATCTTTATCAACTTTTGTAAATTCAAATTTACCTAACTTC	nem316_al2.seq
	TTCTCGTTTTTAATCGTTATTGTAGGATATTCTCTCACATCACCGAATTTT	Majority
	7660 7670 7680 7690 7700	
7520	TTCTCGTTTTTAATCGTTATTGTAGGATATTCTCTCACATCACCGAATTTT	2603_al2.seq
7517	TTCTCGTTTTTAATCGTTATTGTAGGATATTCTCTCACATCACCGAATTTT	nem316_al2.seq
	AGGGATTGGAAAAATCTCTAAGTGTATTAGGATCCTCTGATTTAGGATTCA	Majority
	7710 7720 7730 7740 7750	
7570	AGGGATTGGAAAAATCTCTAAGTGTATTAGGATCCTCTGATTTAGGATTCA	2603_al2.seq
7567	AGGGATTGGAAAAATCTCTAAGTGTATTAGGATCCTCTGATTTAGGATTCA	nem316_al2.seq
	ATGTTGTTCTACCATTAGTGTATAGAAATTTGTTACTTATAAAAAGTGTCA	Majority
	7760 7770 7780 7790 7800	
7620	ATGTTGTTCTACCATTAGTGTATAGAAATTTGTTACTTATAAAAAGTGTCA	2603_al2.seq
7617	ATGTTGTTCTACCATTAGTGTATAGAAATTTGTTACTTATAAAAAGTGTCA	nem316_al2.seq

FIGURE 20K

100/487

	TCTAGTTTTCACATCATATGTGAGTGTTACTTTTTTGTCTTCTCCTAAGTT	Majority
	7810 7820 7830 7840 7850	
7670	TCTAGTTTTCACATCATATGTGAGTGTTACTTTTTTGTCTTCTCCTAAGTT	2603_a12.seq
7667	TCTAGTTTTCACATCATATGTGAGTGTTACTTTTTTGTCTTCTCCTAAGTT	nen316_a12.seq
	CAAAACCTCTAACCTAGAGTTTATTTTTTGTATGTAATTCTAATTTAACCCTT	Majority
	7860 7870 7880 7890 7900	
7720	CAAAACCTCTAACCTAGAGTTTATTTTTTGTATGTAATTCTAATTTAACCCTT	2603_a12.seq
7717	CAAAACCTCTAACCTAGAGTTTATTTTTTGTATGTAATTCTAATTTAACCCTT	nen316_a12.seq
	TAAGTATTTCCACCATCATTATTAGGCCCCACCAGTTGCAATGCTATCTTTT	Majority
	7910 7920 7930 7940 7950	
7770	TAAGTATTTCCACCATCATTATTAGGCCCCACCAGTTGCAATGCTATCTTTT	2603_a12.seq
7767	TAAGTATTTCCACCATCATTATTAGGCCCCACCAGTTGCAATGCTATCTTTT	nen316_a12.seq
	ATTATACTTCCATCATTTCCTGTAAAGTATAATCACTTGCTTGTAATGT	Majority
	7960 7970 7980 7990 8000	
7820	ATTATACTTCCATCATTTCCTGTAAAGTATAATCACTTGCTTGTAATGT	2603_a12.seq
7817	ATTATACTTCCATCATTTCCTGTAAAGTATAATCACTTGCTTGTAATGT	nen316_a12.seq
	TTGTCCGTTGCCAAGCTGTAAATTGATTTTGTCAACCATAGGATCTTCTA	Majority
	8010 8020 8030 8040 8050	
7870	TTGTCCGTTGCCAAGCTGTAAATTGATTTTGTCAACCATAGGATCTTCTA	2603_a12.seq
7867	TTGTCCGTTGCCAAGCTGTAAATTGATTTTGTCAACCATAGGATCTTCTA	nen316_a12.seq
	TAGTTCCATTAACAATTGAGTTTTCTTTTGTAAATCTTTTCAAAATTGT	Majority
	8060 8070 8080 8090 8100	
7920	TAGTTCCATTAACAATTGAGTTTTCTTTTGTAAATCTTTTCAAAATTGT	2603_a12.seq
7917	TAGTTCCATTAACAATTGAGTTTTCTTTTGTAAATCTTTTCAAAATTGT	nen316_a12.seq
	TGCTGAATTTTATAGATAAAATTTTCAATTGTTAGATGTATCGGCTGAAGTTAC	Majority
	8110 8120 8130 8140 8150	
7970	TGCTGAATTTTATAGATAAAATTTTCAATTGTTAGATGTATCGGCTGAAGTTAC	2603_a12.seq
7967	TGCTGAATTTTATAGATAAAATTTTCAATTGTTAGATGTATCGGCTGAAGTTAC	nen316_a12.seq
	TATCGGGGTGTAGTACTCAGGTTTGGAAAGAGAATGACTTCATTAGTTCTG	Majority
	8160 8170 8180 8190 8200	
8020	TATCGGGGTGTAGTACTCAGGTTTGGAAAGAGAATGACTTCATTAGTTCTG	2603_a12.seq
8017	TATCGGGGTGTAGTACTCAGGTTTGGAAAGAGAATGACTTCATTAGTTCTG	nen316_a12.seq
	TTATTTCTCCATCTGAAAGTTTAAAGCTTCTCTTTTCAATTTTTTGAAAA	Majority
	8210 8220 8230 8240 8250	
8070	TTATTTCTCCATCTGAAAGTTTAAAGCTTCTCTTTTCAATTTTTTGAAAA	2603_a12.seq
8067	TTATTTCTCCATCTGAAAGTTTAAAGCTTCTCTTTTCAATTTTTTGAAAA	nen316_a12.seq
	GTACCATCTTGATTTTTTCTTATACTCCTCATTATAAACTTGTCTAAAAAGC	Majority
	8260 8270 8280 8290 8300	
8120	GTACCATCTTGATTTTTTCTTATACTCCTCATTATAAACTTGTCTAAAAAGC	2603_a12.seq
8117	GTACCATCTTGATTTTTTCTTATACTCCTCATTATAAACTTGTCTAAAAAGC	nen316_a12.seq
	AGATATATCTATACCAAAATTAAGATGTCATAATTTTTCTGTTTTTAAAC	Majority
	8310 8320 8330 8340 8350	
8170	AGATATATCTATACCAAAATTAAGATGTCATAATTTTTCTGTTTTTAAAC	2603_a12.seq
8167	AGATATATCTATACCAAAATTAAGATGTCATAATTTTTCTGTTTTTAAAC	nen316_a12.seq
	TATTTATATAAAGTTTGGTTGGTTCATGTTCTTTTACTGGTCCATTT	Majority
	8360 8370 8380 8390 8400	
8220	TATTTATATAAAGTTTGGTTGGTTCATGTTCTTTTACTGGTCCATTT	2603_a12.seq
8217	TATTTATATAAAGTTTGGTTGGTTCATGTTCTTTTACTGGTCCATTT	nen316_a12.seq
	CGATAAAATTGTACCTTTAGGGTAATTAAGATTAAATCTAAATAATGAAG	Majority
	8410 8420 8430 8440 8450	
8270	CGATAAAATTGTACCTTTAGGGTAATTAAGATTAAATCTAAATAATGAAG	2603_a12.seq
8267	CGATAAAATTGTACCTTTAGGGTAATTAAGATTAAATCTAAATAATGAAG	nen316_a12.seq

FIGURE 20L

101/487

TTTTTTC TAAAGT TTTCCAGAGATTATCTGTGTTTGATAACTATCTAAGGGAA Majority
8460 8470 8480 8490 8500

8320 TTTTTC TAAAGT TTTCCAGAGATTATCTGTGTTTGATAACTATCTAAGGGAA 2603_a12.seq
8317 TTTTTC TAAAGT TTTCCAGAGATTATCTGTGTTTGATAACTATCTAAGGGAA nem316_a12.seq

ACAAAAAGTAACTCTCCCCATTTCCITTTTATATCCTCGGGCTTATCAGTA Majority
8510 8520 8530 8540 8550

8370 AAAAAAGTAACTCTCCCCATTTCCITTTTATATCCTCGGGCTTATCAGTA 2603_a12.seq
8367 AAAAAAGTAACTCTCCCCATTTCCITTTTATATCCTCGGGCTTATCAGTA nem316_a12.seq

AGTAGAAAAATTACTTTTATTTAGATATCCATTTTTTTTTTCATTTGTTCAA Majority
8560 8570 8580 8590 8600

8420 AGTAGAAAAATTACTTTTATTTAGATATCCATTTTTTTTTTCATTTGTTCAA 2603_a12.seq
8417 AGTAGAAAAATTACTTTTATTTAGATATCCATTTTTTTTTTCATTTGTTCAA nem316_a12.seq

TTGGCTTTTCATATGATGCACCCAGTTTAAAATTATTAATAGCATATGATC Majority
8610 8620 8630 8640 8650

8470 TTGGCTTTTCATATGATGCACCCAGTTTAAAATTATTAATAGCATATGATC 2603_a12.seq
8467 TTGGCTTTTCATATGATGCACCCAGTTTAAAATTATTAATAGCATATGATC nem316_a12.seq

TTGTTGGAACACCATCAGTTATATGAACAATAATTTTTTGGACTATTTTCA Majority
8660 8670 8680 8690 8700

8520 TTGTTGGAACACCATCAGTTATATGAACAATAATTTTTTGGACTATTTTCA 2603_a12.seq
8517 TTGTTGGAACACCATCAGTTATATGAACAATAATTTTTTGGACTATTTTCA nem316_a12.seq

TTTACTTGACTCAAAATATCATCTGCCCTCCATGAAGGCTTTTCATAGTAA Majority
8710 8720 8730 8740 8750

8570 TTTACTTGACTCAAAATATCATCTGCCCTCCATGAAGGCTTTTCATAGTAA 2603_a12.seq
8567 TTTACTTGACTCAAAATATCATCTGCCCTCCATGAAGGCTTTTCATAGTAA nem316_a12.seq

TGTTTCTCCTACTTTACTAAGATAGTACTGCTTTTGTTCCTCTGGAGTTA Majority
8760 8770 8780 8790 8800

8620 TGTTTCTCCTACTTTACTAAGATAGTACTGCTTTTGTTCCTCTGGAGTTA 2603_a12.seq
8617 TGTTTCTCCTACTTTACTAAGATAGTACTGCTTTTGTTCCTCTGGAGTTA nem316_a12.seq

GTCCGTTTGTAGTTGATCCCCATTTAGCTTTAGGAGCTTCTGTCCGAATC Majority
8810 8820 8830 8840 8850

8670 GTCCGTTTGTAGTTGATCCCCATTTAGCTTTAGGAGCTTCTGTCCGAATC 2603_a12.seq
8667 GTCCGTTTGTAGTTGATCCCCATTTAGCTTTAGGAGCTTCTGTCCGAATC nem316_a12.seq

CTTTTTATAATCTCTTCAGCATTATTTGTTAATTGTTTATGACTATAATT Majority
8860 8870 8880 8890 8900

8720 CTTTTTATAATCTCTTCAGCATTATTTGTTAATTGTTTATGACTATAATT 2603_a12.seq
8717 CTTTTTATAATCTCTTCAGCATTATTTGTTAATTGTTTATGACTATAATT nem316_a12.seq

CTCTGCTCTGAATTGTGAACITTAGTTTGAAGGCCATAATATTTATCATCTT Majority
8910 8920 8930 8940 8950

8770 CTCTGCTCTGAATTGTGAACITTAGTTTGAAGGCCATAATATTTATCATCTT 2603_a12.seq
8767 CTCTGCTCTGAATTGTGAACITTAGTTTGAAGGCCATAATATTTATCATCTT nem316_a12.seq

CTTTAAATCCTTTTACGACATCTACACTCCTACCATCAAAAATATCTGAA Majority
8960 8970 8980 8990 9000

8820 CTTTAAATCCTTTTACGACATCTACACTCCTACCATCAAAAATATCTGAA 2603_a12.seq
8817 CTTTAAATCCTTTTACGACATCTACACTCCTACCATCAAAAATATCTGAA nem316_a12.seq

CCATAGGTAACATAATGCAACCCTATTATCACTGTTTGCTCCTAAAAATATC Majority
9010 9020 9030 9040 9050

8870 CCATAGGTAACATAATGCAACCCTATTATCACTGTTTGCTCCTAAAAATATC 2603_a12.seq
8867 CCATAGGTAACATAATGCAACCCTATTATCACTGTTTGCTCCTAAAAATATC nem316_a12.seq

TTTTACTGCGGTCCCAAGAGCTTTCGGCAGCTTTCTTGGCTTTTATTATGCC Majority
9060 9070 9080 9090 9100

8920 TTTTACTGCGGTCCCAAGAGCTTTCGGCAGCTTTCTTGGCTTTTATTATGCC 2603_a12.seq
8917 TTTTACTGCGGTCCCAAGAGCTTTCGGCAGCTTTCTTGGCTTTTATTATGCC nem316_a12.seq

FIGURE 20M

102/487

	TTTGAAAATTTTGGCCCATCGTTATTCATTGAGTTAGAATTATCGAGTACG	Majority
	9110 9120 9130 9140 9150	
8970	TTTGAAAATTTTGGCCCATCGTTATTCATTGAGTTAGAATTATCGAGTACG	2603_a12.seq
8967	TTTGAAAATTTTGGCCCATCGTTATTCATTGAGTTAGAATTATCGAGTACG	nem316_a12.seq
	AAGACAACATCTAACGGCTTTTGTGTTGTCCTGCTTTTACTATGGTTTT	Majority
	9160 9170 9180 9190 9200	
9020	AAGACAACATCTAACGGCTTTTGTGTTGTCCTGCTTTTACTATGGTTTT	2603_a12.seq
9017	AAGACAACATCTAACGGCTTTTGTGTTGTCCTGCTTTTACTATGGTTTT	nem316_a12.seq
	TCCACTGACAGTTAACTCAATTTTATATTTATTATGAGCTAAATCACCTA	Majority
	9210 9220 9230 9240 9250	
9070	TCCACTGACAGTTAACTCAATTTTATATTTATTATGAGCTAAATCACCTA	2603_a12.seq
9067	TCCACTGACAGTTAACTCAATTTTATATTTATTATGAGCTAAATCACCTA	nem316_a12.seq
	CTTCTGAAAATACGTTTAGATAATGTTCCCTCTGGAATTTCTCTTATATGC	Majority
	9260 9270 9280 9290 9300	
9120	CTTCTGAAAATACGTTTAGATAATGTTCCCTCTGGAATTTCTCTTATATGC	2603_a12.seq
9117	CTTCTGAAAATACGTTTAGATAATGTTCCCTCTGGAATTTCTCTTATATGC	nem316_a12.seq
	TCACCTTCACTTGAATATGCGGTTAACTGCTTTTGCCTCTGACTTTCCATT	Majority
	9310 9320 9330 9340 9350	
9170	TCACCTTCACTTGAATATGCGGTTAACTGCTTTTGCCTCTGACTTTCCATT	2603_a12.seq
9167	TCACCTTCACTTGAATATGCGGTTAACTGCTTTTGCCTCTGACTTTCCATT	nem316_a12.seq
	TGGAACCTGAACCTTTAAACATGCTCAAGTTTATAAGATTCCTTTGTATCTT	Majority
	9360 9370 9380 9390 9400	
9220	TGGAACCTGAACCTTTAAACATGCTCAAGTTTATAAGATTCCTTTGTATCTT	2603_a12.seq
9217	TGGAACCTGAACCTTTAAACATGCTCAAGTTTATAAGATTCCTTTGTATCTT	nem316_a12.seq
	CATAAATTCTCTGCGGGGGATACTGCTTATCTAGTTCTTCGTGATTTTGT	Majority
	9410 9420 9430 9440 9450	
9270	CATAAATTCTCTGCGGGGGATACTGCTTATCTAGTTCTTCGTGATTTTGT	2603_a12.seq
9267	CATAAATTCTCTGCGGGGGATACTGCTTATCTAGTTCTTCGTGATTTTGT	nem316_a12.seq
	CCAATTGTGGAATTTTTTATCACCACTATTTTGTATCGTAGTTTTTCCATT	Majority
	9460 9470 9480 9490 9500	
9320	CCAATTGTGGAATTTTTTATCACCACTATTTTGTATCGTAGTTTTTCCATT	2603_a12.seq
9317	CCAATTGTGGAATTTTTTATCACCACTATTTTGTATCGTAGTTTTTCCATT	nem316_a12.seq
	ACTCTCAACCTTAACTTGCCAAAGTCTGGTTAGTCTTTTTTATAACCTTCGG	Majority
	9510 9520 9530 9540 9550	
9370	ACTCTCAACCTTAACTTGCCAAAGTCTGGTTAGTCTTTTTTATAACCTTCGG	2603_a12.seq
9367	ACTCTCAACCTTAACTTGCCAAAGTCTGGTTAGTCTTTTTTATAACCTTCGG	nem316_a12.seq
	CGCGCTGTTTCTTCTGATAAAGTATAAATCTCCAGGTATGAGATTATCAAAA	Majority
	9560 9570 9580 9590 9600	
9420	CGCGCTGTTTCTTCTGATAAAGTATAAATCTCCAGGTATGAGATTATCAAAA	2603_a12.seq
9417	CGCGCTGTTTCTTCTGATAAAGTATAAATCTCCAGGTATGAGATTATCAAAA	nem316_a12.seq
	GTAGCTTCACTCTGTTAGCTCAGCAGTTACTTTTTCTATTTTACTTTCTCG	Majority
	9610 9620 9630 9640 9650	
9470	GTAGCTTCACTCTGTTAGCTCAGCAGTTACTTTTTCTATTTTACTTTCTCG	2603_a12.seq
9467	GTAGCTTCACTCTGTTAGCTCAGCAGTTACTTTTTCTATTTTACTTTCTCG	nem316_a12.seq
	ATGAGCAGTAGTTTTTTAAACAAAGGTAGCTTTTGAAAGTGGTTTCTTCT	Majority
	9660 9670 9680 9690 9700	
9520	ATGAGCAGTAGTTTTTTAAACAAAGGTAGCTTTTGAAAGTGGTTTCTTCT	2603_a12.seq
9517	ATGAGCAGTAGTTTTTTAAACAAAGGTAGCTTTTGAAAGTGGTTTCTTCT	nem316_a12.seq
	GCTCATCTGTCTTTTTTAACAACCTAATTTTCTTTAGCACCATTTTCCCGT	Majority
	9710 9720 9730 9740 9750	
9570	GCTCATCTGTCTTTTTTAACAACCTAATTTTCTTTAGCACCATTTTCCCGT	2603_a12.seq
9567	GCTCATCTGTCTTTTTTAACAACCTAATTTTCTTTAGCACCATTTTCCCGT	nem316_a12.seq

FIGURE 20N

103/487

```

ACGGTACTTTCCCTAAACATTGCTATTAAGCGGTATTTGCGACAAACA Majority
9760 9770 9780 9790 9800
9620 ACGGTACTTTCCCTAAACATTGCTATTAAGCGGTATTTGCGACAAACA 2603_a12.seq
9617 ACGGTACTTTCCCTAAACATTGCTATTAAGCGGTATTTGCGACAAACA nem316_a12.seq

AAAAAGACTTAAACGTCAATATTTTGTAGAAAATTTTGTGATTTTCTCATT Majority
9810 9820 9830 9840 9850
9670 AAAAAGACTTAAACGTCAATATTTTGTAGAAAATTTTGTGATTTTCTCATT 2603_a12.seq
9667 AAAAAGACTTAAACGTCAATATTTTGTAGAAAATTTTGTGATTTTCTCATT nem316_a12.seq

TACAACTCCTATTGTGCGGAAATGTCGTTTCTAAATCTAAGATCAGATAC Majority
9860 9870 9880 9890 9900
9720 TACAACTCCTATTGTGCGGAAATGTCGTTTCTAAATCTAAGATCAGATAC 2603_a12.seq
9717 TACAACTCCTATTGTGCGGAAATGTCGTTTCTAAATCTAAGATCAGATAC nem316_a12.seq

AGAATATCCTAGAAATATACAACTATCACTTATTATGATATCAATAATTT Majority
9910 9920 9930 9940 9950
9770 AGAATATCCTAGAAATATACAACTATCACTTATTATGATATCAATAATTT 2603_a12.seq
9767 AGAATATCCTAGAAATATACAACTATCACTTATTATGATATCAATAATTT nem316_a12.seq

CTTATTATAAGGTATGGAATTTTAAATGTTTTTCCCAATTTTGAATGAT Majority
9960 9970 9980 9990 10000
9820 CTTATTATAAGGTATGGAATTTTAAATGTTTTTCCCAATTTTGAATGAT 2603_a12.seq
9817 CTTATTATAAGGTATGGAATTTTAAATGTTTTTCCCAATTTTGAATGAT nem316_a12.seq

TTTTCTTTTTATTTGATAATCTTATTTTTTATTATCTTAGAAATATTTCA Majority
10010 10020 10030 10040 10050
9870 TTTTCTTTTTATTTGATAATCTTATTTTTTATTATCTTAGAAATATTTCA 2603_a12.seq
9867 TTTTCTTTTTATTTGATAATCTTATTTTTTATTATCTTAGAAATATTTCA nem316_a12.seq

ATTAGCTTAAGTAGTTGATTTTTCTTTTTTTATGTTTTTAAATATTGCTT Majority
10060 10070 10080 10090 10100
9920 ATTAGCTTAAGTAGTTGATTTTTCTTTTTTTATGTTTTTAAATATTGCTT 2603_a12.seq
9917 ATTAGCTTAAGTAGTTGATTTTTCTTTTTTTATGTTTTTAAATATTGCTT nem316_a12.seq

AAAAATAATGTTTGAGAGAGAGTTTACTGAATTGATTGAAAATTATTTAC Majority
10110 10120 10130 10140 10150
9970 AAAAATAATGTTTGAGAGAGAGTTTACTGAATTGATTGAAAATTATTTAC 2603_a12.seq
9967 AAAAATAATGTTTGAGAGAGAGTTTACTGAATTGATTGAAAATTATTTAC nem316_a12.seq

AAAAAGACATCCTTAATCAAAATAAACTTCTAACTTTTATGCTATGATTAC Majority
10160 10170 10180 10190 10200
10018 AAAAAGACATCCTTAATCAAAATAAACTTCTAACTTTTATGCTATGATTAC 2603_a12.seq
10017 AAAAAGACATCCTTAATCAAAATAAACTTCTAACTTTTATGCTATGATTAC nem316_a12.seq

TACCCCTTCCATTACTCTAGACAAATCATGTCATCAACTTGGTTTATCTGA Majority
10210 10220 10230 10240 10250
10068 TACCCCTTCCATTACTCTAGACAAATCATGTCATCAACTTGGTTTATCTGA 2603_a12.seq
10067 TACCCCTTCCATTACTCTAGACAAATCATGTCATCAACTTGGTTTATCTGA nem316_a12.seq

ACTACTTATTAGGAAATATTGTCATGATTAAACAACCTTTATTTAACAGTC Majority
10260 10270 10280 10290 10300
10118 ACTACTTATTAGGAAATATTGTCATGATTAAACAACCTTTATTTAACAGTC 2603_a12.seq
10117 ACTACTTATTAGGAAATATTGTCATGATTAAACAACCTTTATTTAACAGTC nem316_a12.seq

AACTCTCTCTGAATATCGAAAAGAGTACAATAGTCTACCAATCTAATGGT Majority
10310 10320 10330 10340 10350
10168 AACTCTCTCTGAATATCGAAAAGAGTACAATAGTCTACCAATCTAATGGT 2603_a12.seq
10167 AACTCTCTCTGAATATCGAAAAGAGTACAATAGTCTACCAATCTAATGGT nem316_a12.seq

GTAAGTACAGAAACAAGCTTTCAAATATATTTATCATCAATCACACGTTTT Majority
10360 10370 10380 10390 10400
10218 GTAAGTACAGAAACAAGCTTTCAAATATATTTATCATCAATCACACGTTTT 2603_a12.seq
10217 GTAAGTACAGAAACAAGCTTTCAAATATATTTATCATCAATCACACGTTTT nem316_a12.seq

```

FIGURE 200

104/487

	ACAACTTTTAAATTTTGGATTCAGGCAGGTTACCTTTAA	Majority
	10410 10420 10430 10440 10450	
10268	ACAACTTTTAAATTTTGGATTCAGGCAGGTTACCTTTAA	2603_a12.seq
10267	ACAACTTTTAAATTTTGGATTCAGGCAGGTTACCTTTAA	nen316_a12.seq
	CTTACTTTAGTGAAAAATTTGGACTATCTTGTGCAACTGCTTATCGCAT	Majority
	10460 10470 10480 10490 10500	
10318	CTTACTTTAGTGAAAAATTTGGACTATCTTGTGCAACTGCTTATCGCAT	2603_a12.seq
10317	CTTACTTTAGTGAAAAATTTGGACTATCTTGTGCAACTGCTTATCGCAT	nen316_a12.seq
	CGAAAAACATATTAGTCCGTTACTAGAAAAACCTGGATTTCAGATTTTCAA	Majority
	10510 10520 10530 10540 10550	
10368	CGAAAAACATATTAGTCCGTTACTAGAAAAACCTGGATTTCAGATTTTCAA	2603_a12.seq
10367	CGAAAAACATATTAGTCCGTTACTAGAAAAACCTGGATTTCAGATTTTCAA	nen316_a12.seq
	AAATACTATTACCGGTGACGAGTATCGAATTCGCTATTTAATCGCATTTT	Majority
	10560 10570 10580 10590 10600	
10418	AAATACTATTACCGGTGACGAGTATCGAATTCGCTATTTAATCGCATTTT	2603_a12.seq
10417	AAATACTATTACCGGTGACGAGTATCGAATTCGCTATTTAATCGCATTTT	nen316_a12.seq
	TAAATGCTCGATTGTGATAGAAAGTTTATCCCTTCTTAAGATGGATAAA	Majority
	10610 10620 10630 10640 10650	
10468	TAAATGCTCGATTGTGATAGAAAGTTTATCCCTTCTTAAGATGGATAAA	2603_a12.seq
10467	TAAATGCTCGATTGTGATAGAAAGTTTATCCCTTCTTAAGATGGATAAA	nen316_a12.seq
	TTGCTTATCAAACGATTGTTATTAGAATACTCAACTACTTTTACTGCTTC	Majority
	10660 10670 10680 10690 10700	
10518	TTGCTTATCAAACGATTGTTATTAGAATACTCAACTACTTTTACTGCTTC	2603_a12.seq
10517	TTGCTTATCAAACGATTGTTATTAGAATACTCAACTACTTTTACTGCTTC	nen316_a12.seq
	TCATTACTTCCCAAATACATTTATTTTCTTTGATACATTGTTGTCTCTAT	Majority
	10710 10720 10730 10740 10750	
10568	TCATTACTTCCCAAATACATTTATTTTCTTTGATACATTGTTGTCTCTAT	2603_a12.seq
10567	TCATTACTTCCCAAATACATTTATTTTCTTTGATACATTGTTGTCTCTAT	nen316_a12.seq
	CATGGAAACGTATTAATTATAATGTAGTTGTCCCTTACTCATCCCTTTTC	Majority
	10760 10770 10780 10790 10800	
10618	CATGGAAACGTATTAATTATAATGTAGTTGTCCCTTACTCATCCCTTTTC	2603_a12.seq
10617	CATGGAAACGTATTAATTATAATGTAGTTGTCCCTTACTCATCCCTTTTC	nen316_a12.seq
	ATTGAACTACAAAAATATATTTATCTATGATACATTACAATATTGTGTCAA	Majority
	10810 10820 10830 10840 10850	
10668	ATTGAACTACAAAAATATATTTATCTATGATACATTACAATATTGTGTCAA	2603_a12.seq
10667	ATTGAACTACAAAAATATATTTATCTATGATACATTACAATATTGTGTCAA	nen316_a12.seq
	AAATGTTATTATAGATTCCCTTTAAAAATTAATTTAAAAAAGACGATATAG	Majority
	10860 10870 10880 10890 10900	
10718	AAATGTTATTATAGATTCCCTTTAAAAATTAATTTAAAAAAGACGATATAG	2603_a12.seq
10717	AAATGTTATTATAGATTCCCTTTAAAAATTAATTTAAAAAAGACGATATAG	nen316_a12.seq
	ACTATATTTTCTTGCTTACGTTACTTCCCATAACTCTTTTCTAATCCA	Majority
	10910 10920 10930 10940 10950	
10768	ACTATATTTTCTTGCTTACGTTACTTCCCATAACTCTTTTCTAATCCA	2603_a12.seq
10767	ACTATATTTTCTTGCTTACGTTACTTCCCATAACTCTTTTCTAATCCA	nen316_a12.seq
	AATTGGACTGAGAAGCCTATCGATAATGTAATAGCTATTTTCCGAAATTA	Majority
	10960 10970 10980 10990 11000	
10818	AATTGGACTGAGAAGCCTATCGATAATGTAATAGCTATTTTCCGAAATTA	2603_a12.seq
10817	AATTGGACTGAGAAGCCTATCGATAATGTAATAGCTATTTTCCGAAATTA	nen316_a12.seq
	TCCCAAATTCGAAAAATTTACAGCCACTCAAAGATGCTCTTCCCTTAT	Majority
	11010 11020 11030 11040 11050	
10868	TCCCAAATTCGAAAAATTTACAGCCACTCAAAGATGCTCTTCCCTTAT	2603_a12.seq
10867	TCCCAAATTCGAAAAATTTACAGCCACTCAAAGATGCTCTTCCCTTAT	nen316_a12.seq

FIGURE 20P

105/487

```

CTG GCT C C T A T C A T G A G T T G C T A T C A T T T T T T C C G A Majority
11060 11070 11080 11090 11100
10918 CTG GCT C C T A T C A T G A G T T G C T A T C T T T T T T T C C G A 2603_a12.seq
10917 CTG A C T C C T A T C A T G A G T T G C T A T C A T T T T T T C C G A nem316_a12.seq

A C A T T T A T T T T A G G A T T A A A T C A A T T A A T C C C T G A A A C A A T T T C A T T T C C Majority
11110 11120 11130 11140 11150
10967 A C A T T T A T T T T A G G A T T A A A T C A A T T A A T C C C T G A A A C A A T T T C A T T T C C 2603_a12.seq
10967 A C A T T T A T T T T A G G A T T A A A T C A A T T A A T C C C T G A A A C A A T T T C A T T T C C nem316_a12.seq

T T C A T G G A A C T A T C A T A G A C A T G A T A A A T T A A C T A C T A T T C T C C G A C C G A Majority
11160 11170 11180 11190 11200
11017 T T C A T G G A A C T A T C A T A G A C A T G A T A A A T T A A C T A C T A T T C T C C G A C C G A 2603_a12.seq
11017 T T C A T G G A A C T A T C A T A G A C A T G A T A A A T T A A C T A C T A T T C T C C G A C C G A nem316_a12.seq

T A A T T A C A A A T T G C T T A A G T G A A A T T G G A G A A T A C A C G T T T A A G G A A C A A Majority
11210 11220 11230 11240 11250
11067 T A A T T A C A A A T T G C T T A A G T A A A A T T G G A G A A T A C A C G T T T A A G A A C A A 2603_a12.seq
11067 T A A T T A C A A A T T G C T T A A G T G A A A T T G G A G A A T A C A C G T T T A A G G A A C A A nem316_a12.seq

C A T T T T C T T C C T T T T G T G C T C A T C T A G A A A G A A T T A T C A A A A A T C A T A T Majority
11260 11270 11280 11290 11300
11117 C A T T T T C T T C C T T T T G T A C T C A T C T A G A A A G A A T A T C A A A A A T C A T A T 2603_a12.seq
11117 C A T T T T C T T C C T T T T G T G C T C A C T A G A A A G A A T T A T C A A A A A T C A A T AT nem316_a12.seq

T C C T C C G A T A C A G A T A G C C G T A C T A A C T A C A G A C T T T A T T A A T A A C C A A A Majority
11310 11320 11330 11340 11350
11167 T C C T C C G A T A C A G A T A G C C G T A C T A A C T A C A G A C T T T A T T A A T A A C C A A A 2603_a12.seq
11167 T C C T C C G A T A C A A A T A G C C G T A C T A A C T A C A G A C T T T A T T A A T A A C A A A A nem316_a12.seq

T T T T A A C A G A A T G T T T A T T A C A G A G G T T T T C T T C T A A A C A G A T T C A T T T C Majority
11360 11370 11380 11390 11400
11217 T T T T A A C A G A A T G T T A T T A C A G A G A T T T T C T T C T A A A A G A T T C A T T T C 2603_a12.seq
11217 T T T T A A C A G A A T G T T T A T A C A G A G G T T T T C T T C T A A A C A G A T T C A T T T C nem316_a12.seq

C A C C C T T A C T A T T T A T T A A C T G A T G A T C T T T C C A A T A T T A C T A A T C T T A A Majority
11410 11420 11430 11440 11450
11267 C A C C C T T A C T A T T T A T T A A C T G A T G A T C T T T C C A A T A T T A C T A A T C T T A A 2603_a12.seq
11267 C A C C C T T A C T A T T T A T T A A C T G A T G A T C T T T C C A A T A T T A C T A A T C T T A A nem316_a12.seq

T C C A G A T A T T A T T A T T A C C A A T C C A A A G C T T T C T T C C T T T A T C A A A C A T G Majority
11460 11470 11480 11490 11500
11317 T C C A G A A T T A T T A T T A T T A C C A A T A A A A G C T T T C T C C T T T A T C A A A C A T G 2603_a12.seq
11317 T C C A G A T A T T A T T A T T A T T A C C A A T C C A A A G C T T T C T T C C T T T A T A A A C A T G nem316_a12.seq

A G A T T T C T T C A G A G A G T T T A A T T A C A T A T A T T G A T T G A T T C A T A C T T C A Majority
11510 11520 11530 11540 11550
11367 A G A T T T C T T C A G A G A G T T T A A T T A C A T A T A T T G A T T G A T T A A T A C T T C A 2603_a12.seq
11367 A G A T T T C T T C A G A G A G T T T A A T T A C A T A T A T T G A T T G A T T C A T A C T T C A nem316_a12.seq

G A C C A G A T C A A T C A A A T C C A A G A A A T T A T T T C A T C A A T A C A G G A A G A A A A Majority
11560 11570 11580 11590 11600
11417 G A C C A A A T C A A T C A A A T C C A A A A A A T T A T T T C A T C A A T A C A G A A G A A A A 2603_a12.seq
11417 G A C C A G A T C A A T C A A A T C C A A G A A A T T A T T T C A T C A A T A C A A G A A G A A A A nem316_a12.seq

A T A T T G T A A A C T T T T T T G C A A A A A C T A A T G A A A T A A C T A C T C G T A G C T C C T Majority
11610 11620 11630 11640 11650
11467 A T A T T G T A A A C T T T T T T G C A A A A A C T A A T G A A A T A A C T A A T C G T A G C T C C T 2603_a12.seq
11467 A T A T T G T A A A C T T T T T T G C A A A A A A A A T G A A A T A A C T A C T C G T A G C T C C T nem316_a12.seq

A T A A C T C T T A A A A A T T A A C A T T A A A A A G C T A G A G C A T T G T G T A A T G C T C T Majority
11660 11670 11680 11690 11700
11517 A T A A C T C T T A A A A A T T A A C A T T A A A A A G C T A G A G C A T T G T G T A A T G C T C T 2603_a12.seq
11517 A T A A C T C T T A A A A A T T A A C A T T A A A A A G C T A G A G C A T T G T G T A A T G C T C T nem316_a12.seq

```

FIGURE 20Q

106/487

	A G C T T T T T A A T G T T A A T T T T T T G A A T A A T A T A A T C C A A C T T T T C A A C T G Majority				
	11710	11720	11730	11740	11750
11567	A G C T T T T T A A T G T T A A T T T T T T T T G A A T A A T A T A A T C C A A C T T T T C A A C T G 2603_al2.seq				
11567	A G C T T T T T A A T G T T A A T T T T T T T T G A A T A A T A T A A T C C A A C T T T T C A A C T G nem316_al2.seq				
	T T T T T T C C C A T G T G A A A T G T T C T T T A A T T C T T T T A G C A A T A T T C T G T T G T Majority				
	11760	11770	11780	11790	11800
11617	T T T T T T C C C A T G T G A A A T G T T C T T T A A T T C T T T T A G C A A T A T T C T G T T G T 2603_al2.seq				
11617	T T T T T T C C C A T G T G A A A T G T T C T T T A A T T C T T T T A G C A A T A T T C T G T T G T nem316_al2.seq				
	A G T T T C T C T C T T A A T G C C T T A T C T T T T A C T A A T A A A T C A A G A G A T T C A T G Majority				
	11810	11820	11830	11840	11850
11667	A G T T T C T C T C T T T A A T G C C T T A T C T T T T A C T A A T A A A T C A A G A G A T T C A T G 2603_al2.seq				
11667	A G T T T C T C T C T T T A A T G C C T T A T C T T T T A C T A A T A A A T C A A G A G A T T C A T G nem316_al2.seq				
	G A G T G A C T G A G T A T T T T C T T C C A T G A T T C C T A A C T C A G G G C T A T C A A Majority				
	11860	11870	11880	11890	11900
11717	G A G T G A C T G A G T A T T T T C T T C C A T G A T T C C T A A C T C A G G G C T A T C A A 2603_al2.seq				
11717	G A G T G A C T G A G T A T T T T C T T C C A T G A T T C C T A A C T C A G G G C T A T C A A nem316_al2.seq				
	T A A C T T C A A C T G T T C C A C C G C G A T C T G T T G C A A T A A T A G C A C T T G A A A G T Majority				
	11910	11920	11930	11940	11950
11767	T A A C T T C A A C T G T T C C A C C G C G A T C T G T T G C A A T A A T A G C A C T T G A A A G T 2603_al2.seq				
11767	T A A C T T C A A C T G T T C C A C C G C G A T C T G T T G C A A T A A T A G C A C T T G A A A G T nem316_al2.seq				
	A G A C C A G C T T C T A A A A T A G A G G T T G C T A A T C C C T C T G G A T A C A T T G A A G G Majority				
	11960	11970	11980	11990	12000
11817	A G A C C A G C T T C T A A A A T A G A G G T T G C T A A T C C C T C T G G A T A C A T T G A A G G 2603_al2.seq				
11817	A G A C C A G C T T C T A A A A T A G A G G T T G C T A A T C C C T C T G G A T A C A T T G A A G G nem316_al2.seq				
	G T A A A C A A A G A T A T C A G T C T G T G C C A T T A A A G A C A T A G T C T G T T C A A A G T Majority				
	12010	12020	12030	12040	12050
11867	G T A A A C A A A G A T A T C A G T C T G T G C C A T T A A A G A C A T A G T C T G T T C A A A G T 2603_al2.seq				
11867	G T A A A C A A A G A T A T C A G T C T G T G C C A T T A A A G A C A T A G T C T G T T C A A A G T nem316_al2.seq				
	T T A A T T T C C C C A A A A A G T T A A T C T G T T T G G A C T G A T A T T T C T C T T T C A A A Majority				
	12060	12070	12080	12090	12100
11917	T T A A T T T C C C C A A A A A G T T A A T C T G T T T G G A C T G A T A T T T C T C T T T C A A A 2603_al2.seq				
11917	T T A A T T T C C C C A A A A A G T T A A T C T G T T T G G A C T G A T A T T T C T C T T T C A A A nem316_al2.seq				
	T G T G C T A A T T C A G G T C C G T C T C C T G C A A T C T G T A A A T A A A C A T T T T C A G A Majority				
	12110	12120	12130	12140	12150
11967	T G T G C T A A T T C A G G T C C G T C T C C T G C A A T C T G T A A A T A A A C A T T T T C A G A 2603_al2.seq				
11967	T G T G C T A A T T C A G G T C C G T C T C C T G C A A T C T G T A A A T A A A C A T T T T C A G A nem316_al2.seq				
	G T A C T G T G A C A T C G A A A A T G C T T C T A A G A G C A A T T C A A T G C C T T T T T C T T Majority				
	12160	12170	12180	12190	12200
12017	G T A C T G T G A C A T C G A A A A T G C T T C T A A G A G C A A T T C A A T G C C T T T T T C T T 2603_al2.seq				
12017	G T A C T G T G A C A T C G A A A A T G C T T C T A A G A G C A A T T C A A T G C C T T T T T C T T nem316_al2.seq				
	T A A T A A T T C T A C C A G C A T A A A G T G A T G A A A A T A T C A T C A G C A G A T T T T T C A Majority				
	12210	12220	12230	12240	12250
12067	T A A T A A T T C T A C C A G C A T A A A G T G A T G A A A A T A T C A T C A G C A G A T T T T T C A 2603_al2.seq				
12067	T A A T A A T T C T A C C A G C A T A A A G T G A T G A A A A T A T C A T C A G C A G A T T T T T C A nem316_al2.seq				
	A G G T A A G C C G T A C C A G C A A A A T C A G A G C C T A G A C T T T C A G A T A C C G A A T T Majority				
	12260	12270	12280	12290	12300
12117	A G G T A A G C C G T A C C A G C A A A A T C A G A G C C T A G A C T T T C A G A T A C C G A A T T 2603_al2.seq				
12117	A G G T A A G C C G T A C C A G C A A A A T C A G A G C C T A G A C T T T C A G A T A C C G A A T T nem316_al2.seq				
	A T A A A T A A C T C C T T T A G C T T C T A T A T T A A A A T G T T T T A A C C A T T C A A C G C Majority				
	12310	12320	12330	12340	12350
12167	A T A A A T A A C T C C T T T A G C T T C T A T A T T A A A A T G T T T T A A C C A T T C A A C G C 2603_al2.seq				
12167	A T A A A T A A C T C C T T T A G C T T C T A T A T T A A A A T G T T T T A A C C A T T C A A C G C nem316_al2.seq				

FIGURE 20R

107/487

	TTCTCTTGGAATAGCGGATAAATAATCTGACGATAATGCTTAACACGCGCT	Majority
	12360 12370 12380 12390 12400	
12217	TTCTCTTGGAATACCGCATAAAAATCTGGACGATAATGCTTAACACGCGCT	2603_a12.seq
12217	TTCTCTTGGAATACCGCATAAAAATCTGGACGATAATGCTTAACACGCGCT	nem316_a12.seq
	GTGAGAAGATGTTTCATAGATAGCTCCAAAGAAATCTAAAAAACGATTATT	Majority
	12410 12420 12430 12440 12450	
12267	GTGAGAAGATGTTTCATAGATAGCTCCAAAGAAATCTAAAAAACGATTATT	2603_a12.seq
12267	GTGAGAAGATGTTTCATAGATAGCTCCAAAGAAATCTAAAAAACGATTATT	nem316_a12.seq
	GACAGAAAAATGACTTTGACCCATGGTCTAAAAACAATACTAGGTAATGGT	Majority
	12460 12470 12480 12490 12500	
12317	GACAGAAAAATGACTTTGACCCATGGTCTAAAAACAATACTAGGTAATGGT	2603_a12.seq
12317	GACAGAAAAATGACTTTGACCCATGGTCTAAAAACAATACTAGGTAATGGT	nem316_a12.seq
	GATTCTTTTGCAAAAGATAGCCCTTCTAGCGTTGTTAACTGAAAACGCTGTA	Majority
	12510 12520 12530 12540 12550	
12367	GATTCTTTTGCAAAAGATAGCCCTTCTAGCGTTGTTAACTGAAAACGCTGTA	2603_a12.seq
12367	GATTCTTTTGCAAAAGATAGCCCTTCTAGCGTTGTTAACTGAAAACGCTGTA	nem316_a12.seq
	TTACAAATCACAAAATCAATATTTTTCATCTGAAACATATTTTCATCAGCGT	Majority
	12560 12570 12580 12590 12600	
12417	TTACAAATCACAAAATCAATATTTTTCATCTGAAACATATTTTCATCAGCGT	2603_a12.seq
12417	TTACAAATCACAAAATCAATATTTTTCATCTGAAACATATTTTCATCAGCGT	nem316_a12.seq
	GTGTATTCTCGATTTTTTGTTAATAATAGGATAGCGCTGCTTGACAATAT	Majority
	12610 12620 12630 12640 12650	
12467	GTGTATTCTCGATTTTTTGTTAATAATAGGATAGCGCTGCTTGACAATAT	2603_a12.seq
12467	GTGTATTCTCGATTTTTTGTTAATAATAGGATAGCGCTGCTTGACAATAT	nem316_a12.seq
	TTTTGCTCGGTAAACGGTAAATTTTTTCTACCCCTTGCTCTTCATCTATAATC	Majority
	12660 12670 12680 12690 12700	
12517	TTTTGCTCGGTAAACGGTAAATTTTTTCTACCCCTTGCTCTTCATCTATAATC	2603_a12.seq
12517	TTTTGCTCGGTAAACGGTAAATTTTTTCTACCCCTTGCTCTTCATCTATAATC	nem316_a12.seq
	GGTAAATCACCATGATTAGTTGTTACAATAACAACACGGTAGCCACGCTT	Majority
	12710 12720 12730 12740 12750	
12567	GGTAAATCACCATGATTAGTTGTTACAATAACAACACGGTAGCCACGCTT	2603_a12.seq
12567	GGTAAATCACCATGATTAGTTGTTACAATAACAACACGGTAGCCACGCTT	nem316_a12.seq
	AACCAAATCTGCTGTCATTTTTATCTGTATAACGTTCAATACCTCCGAGGA	Majority
	12760 12770 12780 12790 12800	
12617	AACCAAATCTGCTGTCATTTTTATCTGTATAACGTTCAATACCTCCGAGGA	2603_a12.seq
12617	AACCAAATCTGCTGTCATTTTTATCTGTATAACGTTCAATACCTCCGAGGA	nem316_a12.seq
	AGGGTAGATAATATCCTGAGAAAAACAGCAACTGTTTTTACCTTATTTTCC	Majority
	12810 12820 12830 12840 12850	
12667	AGGGTAGATAATATCCTGAGAAAAACAGCAACTGTTTTTACCTTATTTTCC	2603_a12.seq
12667	AGGGTAGATAATATCCTGAGAAAAACAGCAACTGTTTTTACCTTATTTTCC	nem316_a12.seq
	ATATTTATCCACTTTTCATCAATAAGCCATCTTTTAAGCCTTTAATCATAG	Majority
	12860 12870 12880 12890 12900	
12717	ATATTTATCCACTTTTCATCAATAAGCCATCTTTTAAGCCTTTAATCATAG	2603_a12.seq
12717	ATATTTATCCACTTTTCATCAATAAGCCATCTTTTAAGCCTTTAATCATAG	nem316_a12.seq
	CAACTATTTTTTTTCTCTTTTCTCTCTTCTGCTACCAACAACCTCGAACAAT	Majority
	12910 12920 12930 12940 12950	
12767	CAACTATTTTTTTTCTCTTTTCTCTCTTCTGCTACCAACAACCTCGAACAAT	2603_a12.seq
12767	CAACTATTTTTTTTCTCTTTTCTCTCTTCTGCTACCAACAACCTCGAACAAT	nem316_a12.seq
	TCATTTTCGCATAAAATACTAAATATTTGTATCGCTTCTTCTTACCATATTT	Majority
	12960 12970 12980 12990 13000	
12817	TCATTTTCGCATAAAATACTAAATATTTGTATCGCTTCTTCTTACCATATTT	2603_a12.seq
12817	TCATTTTCGCATAAAATACTAAATATTTGTATCGCTTCTTCTTACCATATTT	nem316_a12.seq

FIGURE 20S

108/487

	TTTTATAAATATAGATCGGCATTGCGCTATTCATGTAATATTTTCGAAATGGTG	Majority
	13010 13020 13030 13040 13050	
12867	TTTTATAAATATAGATCGGCATTGCGCTATTCATGTAATATTTTCGAAATGGTG	2603_a12.seq
12867	TTTTATAAATATAGATCGGCATTGCGCTATTCATGTAATATTTTCGAAATGGTG	nen316_a12.seq
	AATGATTCAATACATGAAAAACATGCCCAAATTTTTTTAACTCGTGAAGAG	Majority
	13060 13070 13080 13090 13100	
12917	AATGATTCAATACATGAAAAACATGCCCAAATTTTTTTAACTCGTGAAGAG	2603_a12.seq
12917	AATGATTCAATACATGAAAAACATGCCCAAATTTTTTTAACTCGTGAAGAG	nen316_a12.seq
	TGTCCAATTTTCGTGTAACAGACCAATAAAATTAACCTGATAAGTCTTATA	Majority
	13110 13120 13130 13140 13150	
12967	TGTCCAATTTTCGTGTAACAGACCAATAAAATTAACCTGATAAGTCTTATA	2603_a12.seq
12967	TGTCCAATTTTCGTGTAACAGACCAATAAAATTAACCTGATAAGTCTTATA	nen316_a12.seq
	TCCCATCTCTGACAGACGATAATTTCATTTTCAGAGTCAACAAAATCAATAA	Majority
	13160 13170 13180 13190 13200	
13017	TCCCATCTCTGACAGACGATAATTTCATTTTCAGAGTCAACAAAATCAATAA	2603_a12.seq
13017	TCCCATCTCTGACAGACGATAATTTCATTTTCAGAGTCAACAAAATCAATAA	nen316_a12.seq
	ACATCTCTTCTGCAAAAGCCAGATGTTTCTTCGAAAACGCTCGTTTTTCATT	Majority
	13210 13220 13230 13240 13250	
13067	ACATCTCTTCTGCAAAAGCCAGATGTTTCTTCGAAAACGCTCGTTTTTCATT	2603_a12.seq
13067	ACATCTCTTCTGCAAAAGCCAGATGTTTCTTCGAAAACGCTCGTTTTTCATT	nen316_a12.seq
	AAAGCAGCCGAAGTAATACACTCTTCAATTTCTTTATAGTCAAATTCCTTG	Majority
	13260 13270 13280 13290 13300	
13117	AAAGCAGCCGAAGTAATACACTCTTCAATTTCTTTATAGTCAAATTCCTTG	2603_a12.seq
13117	AAAGCAGCCGAAGTAATACACTCTTCAATTTCTTTATAGTCAAATTCCTTG	nen316_a12.seq
	CATCACTAAATTTTACGGTTTCATATCTTGATACAAACAAGATAACATAC	Majority
	13310 13320 13330 13340 13350	
13167	CATCACTAAATTTTACGGTTTCATATCTTGATACAAACAAGATAACATAC	2603_a12.seq
13167	CATCACTAAATTTTACGGTTTCATATCTTGATACAAACAAGATAACATAC	nen316_a12.seq
	CGACCTTAGGTAAATGAAGGTAATTTTTCATAATTATCTATCAAATCACCT	Majority
	13360 13370 13380 13390 13400	
13217	CGACCTTAGGTAAATGAAGGTAATTTTTCATAATTATCTATCAAATCACCT	2603_a12.seq
13217	CGACCTTAGGTAAATGAAGGTAATTTTTCATAATTATCTATCAAATCACCT	nen316_a12.seq
	AGGACAACCGAATCTTGATCTAAAGTCAAGAACCAATCAAATTCCTTGTC	Majority
	13410 13420 13430 13440 13450	
13267	AGGACAACCGAATCTTGATCTAAAGTCAAGAACCAATCAAATTCCTTGTC	2603_a12.seq
13267	AGGACAACCGAATCTTGATCTAAAGTCAAGAACCAATCAAATTCCTTGTC	nen316_a12.seq
	TACTGCAAATTTGACCGGATACAGTTCAAAGCATATGCAATCCCTTTATTTT	Majority
	13460 13470 13480 13490 13500	
13317	TACTGCAAATTTGACCGGATACAGTTCAAAGCATATGCAATCCCTTTATTTT	2603_a12.seq
13317	TACTGCAAATTTGACCGGATACAGTTCAAAGCATATGCAATCCCTTTATTTT	nen316_a12.seq
	CTGTTAAATAATCAACAGTTAGGTCGCCCTCTTTCATTATAATCGGCTACT	Majority
	13510 13520 13530 13540 13550	
13367	CTGTTAAATAATCAACAGTTAGGTCGCCCTCTTTCATTATAATCGGCTACT	2603_a12.seq
13367	CTGTTAAATAATCAACAGTTAGGTCGCCCTCTTTCATTATAATCGGCTACT	nen316_a12.seq
	AATTGAGAAATTTCTTCTTATTTTTTCGAGCCATTATCTACCATATAGAT	Majority
	13560 13570 13580 13590 13600	
13417	AATTGAGAAATTTCTTCTTATTTTTTCGAGCCATTATCTACCATATAGAT	2603_a12.seq
13417	AATTGAGAAATTTCTTCTTATTTTTTCGAGCCATTATCTACCATATAGAT	nen316_a12.seq
	GTGGCTTACTTTGAGGATAAAATTGCTCGAATGTTCTGATCTAAGCGTTCAA	Majority
	13610 13620 13630 13640 13650	
13467	GTGGCTTACTTTGAGGATAAAATTGCTCGAATGTTCTGATCTAAGCGTTCAA	2603_a12.seq
13467	GTGGCTTACTTTGAGGATAAAATTGCTCGAATGTTCTGATCTAAGCGTTCAA	nen316_a12.seq

FIGURE 20T

109/487

	T A T T G G G G T T A A A G G T G A C A A T A C C C G C T A A A T A T T T C A T G T T C T A T G C T										Majority
	13660 13670 13680 13690 13700										
13517	T A T T G G G G T T A A A G G T G A C A A T A C C C G C T A A A T A T T T C A T G T T C T A T G C T										2603_a12.seq
13517	T A T T G G G G T T A A A G G T G A C A A T A C C C G C T A A A T A T T T C A T G T T C T A T G C T										nem316_a12.seq
	C T T T T C T A A A A T C T C T A A A T A A C T G A A T G A C T G G T G C T T T T G G T T A T A A A A										Majority
	13710 13720 13730 13740 13750										
13567	C T T T T C T A A A A T C T C T A A A T A A C T G A A T G A C T G G T G C T T T T G G T T A T A A A A										2603_a12.seq
13567	C T T T T C T A A A A T C T C T A A A T A A C T G A A T G A C T G G T G C T T T T G G T T A T A A A A										nem316_a12.seq
	A C G A T A C C G A C A T A G A T A G T T A C T G C T A C T A A A C T T T G A A T G A C A T A A T T										Majority
	13760 13770 13780 13790 13800										
13617	A C G A T A C C G A C A T A G A T A G T T A C T G C T A C T A A A C T T T G A A T G A C A T A A T T										2603_a12.seq
13617	A C G A T A C C G A C A T A G A T A G T T A C T G C T A C T A A A C T T T G A A T G A C A T A A T T										nem316_a12.seq
	T A C C A A T G A T A C T G A C A T T T G A G T A T T G A T A T A A T A G A G T A C A G C T C C A C										Majority
	13810 13820 13830 13840 13850										
13667	T A C C A A T G A T A C T G A C A T T T G A G T A T T G A T A T A A T A G A G T A C A G C T C C A C										2603_a12.seq
13667	T A C C A A T G A T A C T G A C A T T T G A G T A T T G A T A T A A T A G A G T A C A G C T C C A C										nem316_a12.seq
	T A A G A G T A G C A G C A A T T A A A T A G C G C A G C A T T C C T C T T G T T A A T T C T T T A										Majority
	13860 13870 13880 13890 13900										
13717	T A A G A G T A G C A G C A A T T A A A T A G C G C A G C A T T C C T C T T G T T A A T T C T T T A										2603_a12.seq
13717	T A A G A G T A G C A G C A A T T A A A T A G C G C A G C A T T C C T C T T G T T A A T T C T T T A										nem316_a12.seq
	A A A G T A A A T A C A T C T C T T A A A G A G A T A G C T T G A T A T A G G G A G A C A A T A A A										Majority
	13910 13920 13930 13940 13950										
13767	A A A G T A A A T A C A T C T C T T A A A G A G A T A G C T T G A T A T A G G G A G A C A A T A A A										2603_a12.seq
13767	A A A G T A A A T A C A T C T C T T A A A G A G A T A G C T T G A T A T A G G G A G A C A A T A A A										nem316_a12.seq
	T T C A G T A A T A A C T G T A G A G A T A A T A G C T C C C A T A G C A C C T A A A A T T G G T A										Majority
	13960 13970 13980 13990 14000										
13817	T T C A G T A A T A A C T G T A G A G A T A A T A G C T C C C A T A G C A C C T A A A A T T G G T A										2603_a12.seq
13817	T T C A G T A A T A A C T G T A G A G A T A A T A G C T C C C A T A G C A C C T A A A A T T G G T A										nem316_a12.seq
	T T A A A A G T A T A T T A A G C A C A A C A T T T G C C A C A A G T C C A A T A A C T G C A G A C										Majority
	14010 14020 14030 14040 14050										
13867	T T A A A A G T A T A T T A A G C A C A A C A T T T G C C A C A A G T C C A A T A A C T G C A G A C										2603_a12.seq
13867	T T A A A A G T A T A T T A A G C A C A A C A T T T G C C A C A A G T C C A A T A A C T G C A G A C										nem316_a12.seq
	A T T G T G T A A G C T T T T G T A C G T C T T G A A G C C A G T A G A T A C T G T G T C C C T A A										Majority
	14060 14070 14080 14090 14100										
13917	A T T G T G T A A G C T T T T G T A C G T C T T G A A G C C A G T A G A T A C T G T G T C C C T A A										2603_a12.seq
13917	A T T G T G T A A G C T T T T G T A C G T C T T G A A G C C A G T A G A T A C T G T G T C C C T A A										nem316_a12.seq
	A C C G T T A C C A T A A G A A A T G C A A A T G A T C A T C A A A										Majority
	14110 14120 14130										
13967	A C C G T T A C C A T A A G A A A T G C A A A T G A T C A T C A A A										2603_a12.seq
13967	A C C G T T A C C A T A A G A A A T G C A A A T G A T C A T C A A A										nem316_a12.seq

Decoration 'Decoration #1': Shade (with solid black) residues that differ from the Consensus.

FIGURE 20U

110/487

		T C C A C A T C G G T C C A A T T A A C A T A T G A C G T G G C G C A T C A C C A G T A A T T C G G	Majority
		10 20 30 40 50	
1		T C C A C A T C G G T C C A A T T A A C A T A T G A C G T G G C G C A T C A C C A G T A A T T C G G	cohl_a12.seq
1		T C C A C A T C G G T C C A A T T A A C A T A T G A C G T G G C G C A T C A C C A G T A A T T C G G	a909_a12.seq
		T G A A T A A C A A T A T G T T T T G G A A T A A T C T C C A G T T G G T C A C A A A T A A T C G A	Majority
		60 70 80 90 100	
51		T G A A T A A C A A T A T G T T T T G G A A T A A T C T C C A G T T G G T C A C A A A T A A T C G A	cohl_a12.seq
51		T G A A T A A C A A T A T G T T T T G G A A T A A T C T C C A G T T G G T C A C A A A T A A T C G A	a909_a12.seq
		A A T A T A G T C T T C T T G A C T T A A C A A A C G T A A A C G A C C T T C A T G G T A A T C T C	Majority
		110 120 130 140 150	
101		A A T A T A G T C T T C T T G A C T T A A C A A A C G T A A A C G A C C T T C A T G G T A A T C T C	cohl_a12.seq
101		A A T A T A G T C T T C T T G A C T T A A C A A A C G T A A A C G A C C T T C A T G G T A A T C T C	a909_a12.seq
		T C T G C A T T C T T G T A T T A G T C A T A A G A T G C A G A A G G T G T A A T T T T A T A C C C	Majority
		160 170 180 190 200	
151		T C T G C A T T C T T G T A T T A T C A T A A G A T G C A G A A G G T G T A A T T T T A T A C C C	cohl_a12.seq
151		T C T G C A T T C T T G T A T T A G T C A T A A G A T G C A G A A G G T G T A A T T T T A T A C C C	a909_a12.seq
		T G A A T A T C A T T A T C C G T A A C A C A T C G A C G A A C A T T T T C C A C C A T C A T A T C	Majority
		210 220 230 240 250	
201		T G A A T A T C A T T A T C C G T A A C A C A T C G A C G A A C A T T T T C C A C C A T C A T A T C	cohl_a12.seq
201		T G A A T A T C A T T A T C C G T A A C A C A T C G A C G A A C A T T T T C C A C C A T C A T A T C	a909_a12.seq
		A T G T G T C T C C C C T G G G A G A C C A T T T A T T A G G T G A G A A A C G A T T T C T A C T T	Majority
		260 270 280 290 300	
251		A T G T G T C T C C C C T G G G A G A C C A T T T A T T A G G T G A G A A A C G A T T T C T A C T T	cohl_a12.seq
251		A T G T G T C T C C C C T G G G A G A C C A T T T A T T A G G T G A G A A A C G A T T T C T A C T T	a909_a12.seq
		T A G G A G C T A A T T C T C G T A T T C T C T T A A C A G T T T T T T G T A A A G G T C A T A T	Majority
		310 320 330 340 350	
301		T A G G A G C T A A T T C T C G T A T T C T C T T A A C A G T T T T T T G T A A A G G T C A T A T	cohl_a12.seq
301		T A G G A G C T A A T T C T C G T A T T C T C T T A A C A G T T T T T T G T A A A G G T C A T A T	a909_a12.seq
		G A A T G T G C T C T A T T T A T T A A T G C A G A A G T T G C T T C A T A A G T T G T C T G A A G	Majority
		360 370 380 390 400	
351		G A A T G T G C T C T A T T T A T T A A T G C A G A A G T T G C T T C A T A A G T T G T C T G A A G	cohl_a12.seq
351		G A A T G T G C T C T A T T T A T T A A T G C A G A A G T T G C T T C A T A A G T T G T C T G A A G	a909_a12.seq
		G C C T A A T T C T A A A G T C A C A T G C A T T C T T T C A G A A A G T T C A C C G A G A T A G T	Majority
		410 420 430 440 450	
401		G C C T A A T T C T A A A G T C A C A T G C A T T C T T T C A G A A A G T T C A C C G A G A T A G T	cohl_a12.seq
401		G C C T A A T T C T A A A G T C A C A T G C A T T C T T T C A G A A A G T T C A C C G A G A T A G T	a909_a12.seq
		A T A T A G T T T C A T C A G G T A A G C A A T C C G G C C T T G T T C C G A T G T T G A T C C C G	Majority
		460 470 480 490 500	
451		A T A T A G T T T C A T C A G G T A A G C A A T C C G G C C T T G T T C C G A T G T T G A T C C C G	cohl_a12.seq
451		A T A T A G T T T C A T C A G G T A A G C A A T C C G G C C T T G T T C C G A T G T T G A T C C C G	a909_a12.seq
		A T A A C T C C T G G C T C A T T A A T A G C C T G T T C G T A A C G C T C T T T A A T T A T C T C	Majority
		510 520 530 540 550	
501		A T A A C T C C T G G C T C A T T A A T A G C C T G T T C G T A A C G C T C T T T A A T T A T C T C	cohl_a12.seq
501		A T A A C T C C T G G C T C A T T A A T A G C C T G T T C G T A A C G C T C T T T A A T T A T C T C	a909_a12.seq
		T A A C T T A G C A T G G G T A T T G G T A A A A T T T T G A A A A T A G A C T A A G T A T T T A T	Majority
		560 570 580 590 600	
551		T A A C T T A G C A T G G G T A T T G G T A A A A T T T T G A A A A T A G A C T A A G T A T T T A T	cohl_a12.seq
551		T A A C T T A G C A T G G G T A T T G G T A A A A T T T T G A A A A T A G A C T A A G T A T T T A T	a909_a12.seq
		T A A C C T C A G G C C A C T T T T C A T G C A T G A A A T C A A T T T C T T T A T A G A A T T G T	Majority
		610 620 630 640 650	
601		T A A C C T C A G G C C A C T T T T C A T G C A T G A A A T C A A T T T C T T T A T A G A A T T G T	cohl_a12.seq
601		T A A C C T C A G G C C A C T T T T C A T G C A T G A A A T C A A T T T C T T T A T A G A A T T G T	a909_a12.seq

Figure 21

111/487

	TCACGAAATAGGAGCTTCTGGAGCAACTATAGCATCCCCTGAACCAGAAAC	Majority
	660 670 680 690 700	
651	TCACGAAATAGGAGCTTCTGGAGCAACTATAGCATCCCCTGAACCAGAAAC	cohl_al2.seq
651	TCACGAAATAGGAGCTTCTGGAGCAACTATAGCATCCCCTGAACCAGAAAC	a909_al2.seq
	TCTGCAAAAAGTGCACCCCTCCTCTAGCAACTGTTCCATCTCTGTTAGGAC	Majority
	710 720 730 740 750	
701	TCTGCAAAAAGTGCACCCCTCCTCTAGCAACTGTTCCATCTCTGTTAGGAC	cohl_al2.seq
701	TCTGCAAAAAGTGCACCCCTCCTCTAGCAACTGTTCCATCTCTGTTAGGAC	a909_al2.seq
	AGTCAAAAACCAGCATCTATAGGTAATTTAAATATTTTTTCTCCAAAGAGT	Majority
	760 770 780 790 800	
751	AGTCAAAAACCAGCATCTATAGGTAATTTAAATATTTTTTCTCCAAAGAGT	cohl_al2.seq
751	AGTCAAAAACCAGCATCTATAGGTAATTTAAATATTTTTTCTCCAAAGAGT	a909_al2.seq
	TCTCGATAATAATCATTAAATCGCACCATAACGTTTTTTCATAGGATAATT	Majority
	810 820 830 840 850	
801	TCTCGATAATAATCATTAAATCGCACCATAACGTTTTTTCATAGGATAATT	cohl_al2.seq
801	TCTCGATAATAATCATTAAATCGCACCATAACGTTTTTTCATAGGATAATT	a909_al2.seq
	GTATCACAATTTTAACTAAAATAACCTCACTACTACAATAAAACTAAAAA	Majority
	860 870 880 890 900	
851	GTATCACAATTTTAACTAAAATAACCTCACTACTACAATAAAACTAAAAA	cohl_al2.seq
851	GTATCACAATTTTAACTAAAATAACCTCACTACTACAATAAAACTAAAAA	a909_al2.seq
	AGATTGGAACGTCAGTTAGTTCCAATCTTTTATTTACTTCACTTTCTTTA	Majority
	910 920 930 940 950	
901	AGATTGGAACGTCAGTTAGTTCCAATCTTTTATTTACTTCACTTTCTTTA	cohl_al2.seq
901	AGATTGGAACGTCAGTTAGTTCCAATCTTTTATTTACTTCACTTTCTTTA	a909_al2.seq
	ACCAATCCTTTGGCTAAAAAGATATACGCCAGTTAGATTCAAAATACCATAA	Majority
	960 970 980 990 1000	
951	ACCAATCCTTTGGCTAAAAAGATATACGCCAGTTAGATTCAAAATACCATAA	cohl_al2.seq
951	ACCAATCCTTTGGCTAAAAAGATATACGCCAGTTAGATTCAAAATACCATAA	a909_al2.seq
	GCAAGTATAAAAACCAGCTAAAACATCTGTCGGAAAAATGAACCCCTAGGTA	Majority
	1010 1020 1030 1040 1050	
1001	GCAAGTATAAAAACCAGCTAAAACATCTGTCGGAAAAATGAACCCCTAGGTA	cohl_al2.seq
1001	GCAAGTATAAAAACCAGCTAAAACATCTGTCGGAAAAATGAACCCCTAGGTA	a909_al2.seq
	AATACGAGATAACCCAATTAAAAAAATGAGCAAACCCAATGTACCTTGGC	Majority
	1060 1070 1080 1090 1100	
1051	AATACGAGATAACCCAATTAAAAAAATGAGCAAACCCAATGTACCTTGGC	cohl_al2.seq
1051	AATACGAGATAACCCAATTAAAAAAATGAGCAAACCCAATGTACCTTGGC	a909_al2.seq
	ACAACAGTTTCCATATACTCTTAGGCATATAGTACTGCAATAAAATAATA	Majority
	1110 1120 1130 1140 1150	
1101	ACAACAGTTTCCATATACTCTTAGGCATATAGTACTGCAATAAAATAATA	cohl_al2.seq
1101	ACAACAGTTTCCATATACTCTTAGGCATATAGTACTGCAATAAAATAATA	a909_al2.seq
	CTACTCCCAAATATCATATAATGTTCCCATCGAGTCCCACTGGCAAACGA	Majority
	1160 1170 1180 1190 1200	
1151	CTACTCCCAAATATCATATAATGTTCCCATCGAGTCCCACTGGCAAACGA	cohl_al2.seq
1151	CTACTCCCAAATATCATATAATGTTCCCATCGAGTCCCACTGGCAAACGA	a909_al2.seq
	ATAGCCACCTGCAAATACTAAATGGGTTAAAGTTGGTCTCACTCTTTGAA	Majority
	1210 1220 1230 1240 1250	
1201	ATAGCCACCTGCAAATACTAAATGGGTTAAAGTTGGTCTCACTCTTTGAA	cohl_al2.seq
1201	ATAGCCACCTGCAAATACTAAATGGGTTAAAGTTGGTCTCACTCTTTGAA	a909_al2.seq
	AAATAAGTTTTTAAAGAAAGTATACATATACCAGAGATAATAGCATTCTACT	Majority
	1260 1270 1280 1290 1300	
1251	AAATAAGTTTTTAAAGAAAGTATACATATACCAGAGATAATAGCATTCTACT	cohl_al2.seq
1251	AAATAAGTTTTTAAAGAAAGTATACATATACCAGAGATAATAGCATTCTACT	a909_al2.seq

FIGURE 21A

112/487

		C C G A T A A A T C T A G C T T G A G G A T A C C A C T T C T T A A G G T A A C A G A A A G T G A C Majority									
		1310	1320	1330	1340	1350					
1301		G C G A T A A A T C T A G C T T G A G G A T A C C A C T T C T T A A G G T A A C A G A A A G T G A C coh1_a12.seq									
1301		G C G A T A A A T C T A G C T T G A G G A T A C C A C T T C T T A A G G T A A C A G A A A G T G A C a909_a12.seq									
		G C T C A T A A T C G C A A T A G C T A T C T G G C T T A C A G T A T T A C C A A C C A C A G T G A Majority									
		1360	1370	1380	1390	1400					
1351		G C T C A T A A T C G C A A T A G C T A T C T G G C T T A C A G T A T T A C C A A C C A C A G T G A coh1_a12.seq									
1351		G C T C A T A A T C G C A A T A G C T A T C T G G C T T A C A G T A T T A C C A A C C A C A G T G A a909_a12.seq									
		T T A A C T T G A A A A A T C T T G T A G A A A G A T T T G G C A A C T G T C C T C T A A C A C T T Majority									
		1410	1420	1430	1440	1450					
1401		T T A A C T T A A A A A A T C T T G T A G A A A G A T T T G G C A A C T G T C C T C T A A C A C T T coh1_a12.seq									
1401		T T A A C T T G A A A A A T C T T G T A G A A A G A T T T G G C A A C T G T C C T C T A A C A C T T a909_a12.seq									
		T C T T G A A T A G T T T G G T C A A A T G C G A T T A C A G T G T C G G G C C A A T A T T T G A T Majority									
		1460	1470	1480	1490	1500					
1451		T C T T G A A T A G T T T G G T C A A A T G C G A T T A C A G T G T C G G G C C A A T A T T T G A T coh1_a12.seq									
1451		T C T T G A A T A G T T T G G T C A A A T G C G A T T A C A G T G T C G G G C C A A T A T T T G A T a909_a12.seq									
		G A C C A A T C C T A A A C T G A A A A A T A A G A T A A T A G C A A T A A A T G C T T G A A T A A Majority									
		1510	1520	1530	1540	1550					
1501		G A C C A A T C C T A A A C T G A A A A A T A A G A T A A T A G C A A T A A A T G C T T G A A T A A coh1_a12.seq									
1501		G A C C A A T C C T A A A C T G A A A A A T A A G A T A A T A G C A A T A A A T G C T T G A A T A A a909_a12.seq									
		G T T T A C T A T T T T G A C G A G A T A A C A T T A G T C T T T T T A T A T C T T T C T A A T A T Majority									
		1560	1570	1580	1590	1600					
1551		G T T T A C T A T T T T G A C G A G A T A A C A T T A G T C T T T T T A T A T C T T T C T A A T A T coh1_a12.seq									
1551		G T T T A C T A T T T T G A C G A G A T A A C A T T A G T C T T T T T A T A T C T T T C T A A T A T a909_a12.seq									
		T G G C A A A C A A G C C A C G T A A G T T A G A T A G A A A A C A A T C G A A A T T A A A A T T C Majority									
		1610	1620	1630	1640	1650					
1601		T G G C A A A C A A G C C A C G T A A G T T A G A T A G A A A A C A A T C G A A A T T A A A A T T C coh1_a12.seq									
1601		T G G C A A A C A A G C C A C G T A A G T T A G A T A G A A A A C A A T C G A A A T T A A A A T T C a909_a12.seq									
		C C T C A A C G A T A T T A A A T G G A A T A A C C A T T G T T A A A A G G T A A T T G C C T A C A Majority									
		1660	1670	1680	1690	1700					
1651		C C T C A A C G A T A T T A A A T G G A A T A A C C A T T G T T A A A A G G T A A T T G C C T A C A coh1_a12.seq									
1651		C C T C A A C G A T A T T A A A T G G A A T A A C C A T T G T T A A A A G G T A A T T G C C T A C A a909_a12.seq									
		C C A A T A A A T G T T C T G A T A T C A A A G T T A G C A A A T A T A G C A T A C A A A G G A A T Majority									
		1710	1720	1730	1740	1750					
1701		C C A A T A A A T G T T C T G A T A T C A A A G T T A G C A A A T A T A G C A T A C A A A G G A A T coh1_a12.seq									
1701		C C A A T A A A T G T T C T G A T A T C A A A G T T A G C A A A T A T A G C A T A C A A A G G A A T a909_a12.seq									
		C G C A A A G A C A T A G T T G A G A G C T A C C A T A G A T A C A G T C A A G C T A A C T G T A C Majority									
		1760	1770	1780	1790	1800					
1751		C G C A A A G A C A T A G T T G A G A G C T A C C A T A G A T A C A G T C A A G C T A A C T G T A C coh1_a12.seq									
1751		C G C A A A G A C A T A G T T G A G A G C T A C C A T									

FIGURE 21B

113/487

	1960	1970	1980	1990	2000	Majority
	GTGATCTTAAATAAGAGTACACCATTAATGATTTCAAATCAAATAAAATA					
1951	GTGATCTTAAATAAGAGTACACCATAACTTGATTTCAAATCAAATAAAATA					cohl_a12.seq
1951	GTGATCTTAAATAAGAGTACACCATAACTTGATTTCAAATCAAATAAAATA					a909_a12.seq
	2010	2020	2030	2040	2050	Majority
	AAAGCAACTAACATCGGAAGGATTGAAAAATCAACCTTTAAAAATTCTGC					
2001	AAAGCAACTAACATCGGAAGGATTGAAAAATCAACCTTTAAAAATTCTGC					cohl_a12.seq
2001	AAAGCAACTAACATCGGAAGGATTGAAAAATCAACCTTTAAAAATTCTGC					a909_a12.seq
	2060	2070	2080	2090	2100	Majority
	TCCTGGTATTAATGGAATGAAACCATCATCAATACAAAAGATAAGGCAG					
2051	TCCTGGTATTAATGGAATGAAACCATCATCAATACAAAAGATAAGGCAG					cohl_a12.seq
2051	TCCTGGTATTAATGGAATGAAACCATCATCAATACAAAAGATAAGGCAG					a909_a12.seq
	2110	2120	2130	2140	2150	Majority
	AAAGAATGGCGATTGTCACCATTTTACGTGTATTTGTCATAAAAAAATTC					
2101	AAAGAATGGCGATTGTCACCATTTTACGTGTATTTGTCATAAAAAAATTC					cohl_a12.seq
2101	AAAGAATGGCGATTGTCACCATTTTACGTGTATTTGTCATAAAAAAATTC					a909_a12.seq
	2160	2170	2180	2190	2200	Majority
	CTCCAATTTAAATAAATTGAAAGAAGCTCCAAAGGTAAGCGTAGGTACGC					
2151	CTCCAATTTAAATAAATTGAAAGAAGCTCCAAAGGTAAGCGTAGGTACGC					cohl_a12.seq
2151	CTCCAATTTAAATAAATTGAAAGAAGCTCCAAAGGTAAGCGTAGGTACGC					a909_a12.seq
	2210	2220	2230	2240	2250	Majority
	GAAAAAAACCTTTTCTCTTCTCCCATCCAGACTTTTACTGTCTGGTTGTGGAA					
2201	GAAAAAAACCTTTTCTCTTCTCCCATCCAGACTTTTACTGTCTGGTTGTGGAA					cohl_a12.seq
2201	GAAAAAAACCTTTTCTCTTCTCCCATCCAGACTTTTACTGTCTGGTTGTGGAA					a909_a12.seq
	2260	2270	2280	2290	2300	Majority
	TCTCACCACATCAGCTTTTCGCTCGCGGACTGATGCTTCACAACCTGACAAA					
2251	TCTCACCACATCAGCTTTTCGCTCGCGGACTGATGCTTCACAACCTGACAAA					cohl_a12.seq
2251	TCTCACCACATCAGCTTTTCGCTCGCGGACTGATGCTTCACAACCTGACAAA					a909_a12.seq
	2310	2320	2330	2340	2350	Majority
	TAAGTTGGAAGCGATTACCGCCGCTCGGGAATTACACCCTGCCCTGAAGA					
2301	TAAGTTGGAAGCGATTACCGCCGCTCGGGAATTACACCCTGCCCTGAAGA					cohl_a12.seq
2301	TAAGTTGGAAGCGATTACCGCCGCTCGGGAATTACACCCTGCCCTGAAGA					a909_a12.seq
	2360	2370	2380	2390	2400	Majority
	CACCTATAGCATAACAAAAAAACTTGC AATTGCAAGTTTTTTAATCACT					
2351	CACCTATAGCATAACAAAAAAACTTGC AATTGCAAGTTTTTTAATCACT					cohl_a12.seq
2351	CACCTATAGCATAACAAAAAAACTTGC AATTGCAAGTTTTTTAATCACT					a909_a12.seq
	2410	2420	2430	2440	2450	Majority
	AATTAGTAGTAGATTGTAATAATATTAATTTTTAACATCAATTAATTGACA					
2401	AATTAGTAGTAGATTGTAATAATATTAATTTTTAACATCAATTAATTGACA					cohl_a12.seq
2401	AATTAGTAGTAGATTGTAATAATATTAATTTTTAACATCAATTAATTGACA					a909_a12.seq
	2460	2470	2480	2490	2500	Majority
	GGGCACTAATACTCTAGCTACTCCTGCCCTTTGTACAAAGTAAACAAGCTTA					
2451	GGGCACTAATACTCTAGCTACTCCTGCCCTTTGTACAAAGTAAACAAGCTTA					cohl_a12.seq
2451	GGGCACTAATACTCTAGCTACTCCTGCCCTTTGTACAAAGTAAACAAGCTTA					a909_a12.seq
	2510	2520	2530	2540	2550	Majority
	AGTCCCAATCATTGCTCTGATGTGGCAGTTTTATAAACTTTTTCAATCGCT					
2501	AGTCCCAATCATTGCTCTGATGTGGCAGTTTTATAAACTTTTTCAATCGCT					cohl_a12.seq
2501	AGTCCCAATCATTGCTCTGATGTGGCAGTTTTATAAACTTTTTCAATCGCT					a909_a12.seq
	2560	2570	2580	2590	2600	Majority
	GTTGGTTCAATAAATTTCTCTATTACTGATTTTGTAGTGATAGATTTGCC					
2551	GTTGGTTCAATAAATTTCTCTATTACTGATTTTGTAGTGATAGATTTGCC					cohl_a12.seq
2551	GTTGGTTCAATAAATTTCTCTATTACTGATTTTGTAGTGATAGATTTGCC					a909_a12.seq

FIGURE 21C

FIGURE 21D

115/487

		GACATTAGCTTCAATAAAGCTTGAGCTTTAGTCTGAATAAATAGATAAAAT	Majority		
	3260	3270	3280	3290	3300
3251		GACATTAGCTTCAATAAAGCTTGAGCTTTAGTCTGAATAAATAGATAAAAT	cohl_a12.seq		
3251		GACATTAGCTTCAATAAAGCTTGAGCTTTAGTCTGAATAAATAGATAAAAT	a909_a12.seq		
		CCCTTGAGGAAAGATTGTTTCGCAACAATAACCTTCAGCCGGTAAATTATCAA	Majority		
	3310	3320	3330	3340	3350
3301		CCCTTGAGGAAAGATTGTTTCGCAACAATAACCTTCAGCCGGTAAATTATCAA	cohl_a12.seq		
3301		CCCTTGAGGAAAGATTGTTTCGCAACAATAACCTTCAGCCGGTAAATTATCAA	a909_a12.seq		
		ACGTTTGATAAAGGTTGAGTTTTATGAACAGCTTTTGTTAGTAGATTGACG	Majority		
	3360	3370	3380	3390	3400
3351		ACGTTTGATAAAGGTTGAGTTTTATGAACAGCTTTTGTTAGTAGATTGACG	cohl_a12.seq		
3351		ACGTTTGATAAAGGTTGAGTTTTATGAACAGCTTTTGTTAGTAGATTGACG	a909_a12.seq		
		TATTTGGCTTGGTTACTATCAAGCTTTACTTGTGTTAGATCATCGTCTTT	Majority		
	3410	3420	3430	3440	3450
3401		TATTTGGCTTGGTTACTATCAAGCTTTACTTGTGTTAGATCATCGTCTTT	cohl_a12.seq		
3401		TATTTGGCTTGGTTACTATCAAGCTTTACTTGTGTTAGATCATCGTCTTT	a909_a12.seq		
		TATTCGAATACCTTGAAATGGGGCTAGTTAGAGTAAAAACTTGGTTACCAT	Majority		
	3460	3470	3480	3490	3500
3451		TATTCGAATACCTTGAAATGGGGCTAGTTAGAGTAAAAACTTGGTTACCAT	cohl_a12.seq		
3451		TATTCGAATACCTTGAAATGGGGCTAGTTAGAGTAAAAACTTGGTTACCAT	a909_a12.seq		
		GAACATCTTTAGCTTGTGCTACTTGGTAAACAAGTAAATTACCGCCAGCG	Majority		
	3510	3520	3530	3540	3550
3501		GAACATCTTTAGCTTGTGCTACTTGGTAAACAAGTAAATTACCGCCAGCG	cohl_a12.seq		
3501		GAACATCTTTAGCTTGTGCTACTTGGTAAACAAGTAAATTACCGCCAGCG	a909_a12.seq		
		ATACCTTGATTATTATACCTTATTTTGTATAGTAATAGAACCCTTTTCAT	Majority		
	3560	3570	3580	3590	3600
3551		ATACCTTGATTATTATACCTTATTTTGTATAGTAATAGAACCCTTTTCAT	cohl_a12.seq		
3551		ATACCTTGATTATTATACCTTATTTTGTATAGTAATAGAACCCTTTTCAT	a909_a12.seq		
		CTGATCATTGGTATCAGCAGACACAAGTTGAGTACTTAGACTAAATAATA	Majority		
	3610	3620	3630	3640	3650
3601		CTGATCATTGGTATCAGCAGACACAAGTTGAGTACTTAGACTAAATAATA	cohl_a12.seq		
3601		CTGATCATTGGTATCAGCAGACACAAGTTGAGTACTTAGACTAAATAATA	a909_a12.seq		
		AGAGAAGAGTTATCTTTAGGATCTTTTTATAAATCATTGTTCTCTTCCTT	Majority		
	3660	3670	3680	3690	3700
3651		AGAGAAGAGTTATCTTTAGGATCTTTTTATAAATCATTGTTCTCTTCCTT	cohl_a12.seq		
3651		AGAGAAGAGTTATCTTTAGGATCTTTTTATAAATCATTGTTCTCTTCCTT	a909_a12.seq		
		TCTCATTGCTTGTTTTAAATTTTCTTACGTTGACGTGCTCTCCTAGTTA	Majority		
	3710	3720	3730	3740	3750
3701		TCTCATTGCTTGTTTTAAATTTTCTTACGTTGACGTGCTCTCCTAGTTA	cohl_a12.seq		
3701		TCTCATTGCTTGTTTTAAATTTTCTTACGTTGACGTGCTCTCCTAGTTA	a909_a12.seq		
		CTTCTAAAGAGATTAAAAAGTAAAAATCAAAGTAAGGAAAAATAGCGATAAAAT	Majority		
	3760	3770	3780	3790	3800
3751		CTTCTAAAGAGATTAAAAAGTAAAAATCAAAGTAAGGAAAAATAGCGATAAAAT	cohl_a12.seq		
3751		CTTCTAAAGAGATTAAAAAGTAAAAATCAAAGTAAGGAAAAATAGCGATAAAAT	a909_a12.seq		
		GGTGGGATATAAAATAGGCTCTATTGTATTGCCCTCTGCTACTACCAAAGC	Majority		
	3810	3820	3830	3840	3850
3801		GGTGGGATATAAAATAGGCTCTATTGTATTGCCCTCTGCTACTACCAAAGC	cohl_a12.seq		
3801		GGTGGGATATAAAATAGGCTCTATTGTATTGCCCTCTGCTACTACCAAAGC	a909_a12.seq		
		GTTACCAATTATCGTTTGGTACACGATGTCCTCTCACTAGTAACCGATGGG	Majority		
	3860	3870	3880	3890	3900
3851		GTTACCAATTATCGTTTGGTACACGATGTCCTCTCACTAGTAACCGATGGG	cohl_a12.seq		
3851		GTTACCAATTATCGTTTGGTACACGATGTCCTCTCACTAGTAACCGATGGG	a909_a12.seq		

FIGURE 21E

116/487

	T A T T A A C G C C A T A T G G T G T A C A C G T C A C C A A A G T T T G G T A G T C T T T A C C T	Majority
	3910 3920 3930 3940 3950	
3901	T A T T A A C G C C A T A T G G T G T A C A C G T C A C C A A A G T T T G G T A G T C T T T A C C T	cohl_al2.seq
3901	T A T T A A C A C C A T A T G G T G T A C A C G T C A C C A A A G T T T G G T A G T C T T T A C C T	a909_al2.seq
	T T A A C A A T T T G T A A A T C C C T C A A A T C A T C C G G T T T A A C T G T T C T G A T T T G	Majority
	3960 3970 3980 3990 4000	
3951	T T A A C A A T T T G T A A A T C C C T C A A A T C A T C C G G T T T A A C T G T T C T G A T T T G	cohl_al2.seq
3951	T T A A C A A T T T G T A A A T C C C T C A A A T C A T C C G G T T T A A C T G T T C T G A T T T G	a909_al2.seq
	A T C C A C T T G A T A A G T A T A T G T T T C A T T T A A G A T A C T G A C T G T C C A G T G G T	Majority
	4010 4020 4030 4040 4050	
4001	A T C C A C T T G A T A A G T A T A T G T T T C A T T T A A G A T A C T G A C T G T C C A G T G G T	cohl_al2.seq
4001	A T C C A C T T G A T A A G T A T A T G T T T C A T T T A A G A T A C T G A C T G T C C A G T G G T	a909_al2.seq
	C T C C A G C T T T T A A C T T A T C C A A A T C A G A A A A A G C C T T G A A G A G G G T A A A	Majority
	4060 4070 4080 4090 4100	
4051	C T C C A A C T T T T A A C T T A T C C A A A T C A G A A A A A G C C T T G A A G A G G G T A A A	cohl_al2.seq
4051	C T C C A G C T T T T A A C T T A T C C A A A T C A G A A A A A G C C T T G A A G A G G G T A A A	a909_al2.seq
	C C T C T A T G T C C T G A T A A A A T A G A A T G A G T T G A G T C T C C T C C A A T T G G A A G	Majority
	4110 4120 4130 4140 4150	
4101	C C T C T A T G T C C T G A T A A A A T A G A A T G A G T T G A G T C T C C T C C A A T T G G A A G	cohl_al2.seq
4101	C C T C T A T G T C C T G A T A A A A T A G A A T G A G T T G A G T C T C C T C C A A T T G G A A G	a909_al2.seq
	A C T A C T T C C T T C T A A A T G A C C A A T A G A A G T T T G A A G C A C T T T T T C A C T T G	Majority
	4160 4170 4180 4190 4200	
4151	A C T A C T T C C T T C T A A A T G A C C A A T A G A A G T T T G A A G C A C T T T T T C A C T T G	cohl_al2.seq
4151	A C T A C T T C C T T C T A A A T G A C C A A T A G A A G T T T G A A G C A C T T T T T C A C T T G	a909_al2.seq
	T A C C A T G A T A A A G T G G T A A T T T T A T G T T T A T C T T T G C A A T T G A A A T A T A A	Majority
	4210 4220 4230 4240 4250	
4201	T A C C A T G A T A A A G T G G T A A T T T T A T G T T T A T C T T T G C A A T T G A A A T A T A A	cohl_al2.seq
4201	T A C C A T G A T A A A G T G G T A A T T T T A T G T T T A T C T T T G C A A T T G A A A T A T A A	a909_al2.seq
	C C C A T A T T A C C C G T T T T A T C G A T A G C C A G T T G T G A A T T A T A A T C C A A A C G	Majority
	4260 4270 4280 4290 4300	
4251	C C C A T A T T A C C C G T T T T A T C G A T A G C C A G T T G T G A A T T A T A A T C C A A A C G	cohl_al2.seq
4251	C C C A T A T T A C C C G T T T T A T C G A T A G C C A G T T G T G A A T T A T A A T C C A A A C G	a909_al2.seq
	C T C T E G G T T A G T C A T G T G C C A C T T T C A T T C C T G A A G T T T T A A A T T G C T T A T	Majority
	4310 4320 4330 4340 4350	
4301	C T C T T G G T T A G T C A T G T G C C A C T T T C A T T C C T G A A G T T T T A A A T T G C T T A T	cohl_al2.seq
4301	C T C T T G G T T A G T C A T G T G C C A C T T T C A T T C C T G A A G T T T T A A A T T G C T T A T	a909_al2.seq
	T A T A T T C T T T G G C T C G G T T A A T A A T T T T T T A T A G T C G T T T T C A T C C A T A	Majority
	4360 4370 4380 4390 4400	
4351	T A T A T T C T T T G G C T C G G T T A A T A A T T T T T T A T A T C G T T T T C A T C C A T A	cohl_al2.seq
4351	T A T A T T C T T T G G C T C G G T T A A T A A T T T T T T A T A T C G T T T T C A T C C A T A	a909_al2.seq
	T G C G T T A C C G G T C T T G G T A A T C G A T A A T C G C T C G A G A T T G G T G A A A T G A	Majority
	4410 4420 4430 4440 4450	
4401	T G C G T T A C C G G T C T T G G T A A T C G A T A A T C G C T C G A G A T T G G T G A A A T G A	cohl_al2.seq
4401	T G C G T T A C C G G T C T T G G T A A T C G A T A A T C G C T C G A G A T T G G T G A A A T G A	a909_al2.seq
	A T T C C A A T A A T A G C A A G T G A A G G A T A A G C C A T T A A G C C T A C C C C A C T G	Majority
	4460 4470 4480 4490 4500	
4451	A T T C C A A T A A T A G C A A G T G A A G G A T A A G C C A T T A A G C C T A C C C C A C T G	cohl_al2.seq
4451	A T T C C A A T A A T A G C A A G T G A A G G A T A A G C C A T T A A G C C T A C C C C A C T G	a909_al2.seq
	C A A T T A T A G T G A C A A G C A A A A T G G A T A C T A A A T G T T G T C T T A T T T T T T T C	Majority
	4510 4520 4530 4540 4550	
4501	C A A T T A T A G T G A C A A G C A A A A T G G A T A C T A A A T G T T G T C T T A T T T T T T T C	cohl_al2.seq
4501	C A A T T A T A G T G A C A A G C A A A A T G G A T A C T A A A T G T T G T C T T A T T T T T T T C	a909_al2.seq

FIGURE 21F

117/487

```

ATATATTTTAAAGCTGTAACCGCTTCTAGCCCCATCTTATTAAGAACGTA Majority
4560 4570 4580 4590 4600
4551 ATATATTTTAAATCTGTACCACCTTTGCTAGCCCCATCTTATTAAGAACGTA cohl_al2.seq
4551 ATATATTTTAAATCTGTACCACCTTTGCTAGCCCCATCTTATTAAGAACGTA a909_al2.seq

AACGACGACGAGCAACAAGCACGATACCTGCTCCTATTACTAAAATTGCA Majority
4610 4620 4630 4640 4650
4601 AACGACGACGAGCAACAAGCACGATACCTGCTCCTATTACTAAAATTGCA cohl_al2.seq
4601 AACGACGACGAGCAACAAGCACGATACCTGCTCCTATTACTAAAATTGCA a909_al2.seq

CCTATAATGTAGAAAATTGTTGTACCAATACCACCTGTTGAAGGCAACTC Majority
4660 4670 4680 4690 4700
4651 CCTATAATGTAGAAAATTGTTGTACCAATACCACCTGTTGAAGGCAACTC cohl_al2.seq
4651 CCTATAATGTAGAAAATTGTTGTACCAATACCACCTGTTGAAGGCAACTC a909_al2.seq

AGTACCTTTGTTATTTTCAACAGTTGGGTTAACTAAAAGGTTATCTGAAT Majority
4710 4720 4730 4740 4750
4701 AGTACCTTTGTTATTTTCAACAGTTGGGTTAACTAAAAGGTTATCTGAAT cohl_al2.seq
4701 AGTACCTTTGTTATTTTCAACAGTTGGGTTAACTAAAAGGTTATCTGAAT a909_al2.seq

TAGTCGTATCAGTGGCTCCATCTCCTAAAATAACCTTCTGAGAGTTATCT Majority
4760 4770 4780 4790 4800
4751 TAGTCGTATCAGTGGCTCCATCTCCTAAAATAACCTTCTGAGAGTTATCT cohl_al2.seq
4751 TAGTCGTATCAGTGGCTCCATCTCCTAAAATAACCTTCTGAGAGTTATCT a909_al2.seq

AACAAATTGTAACTAAGGGAGCCTTTTTCTCAACTAGATAGTATGTACC Majority
4810 4820 4830 4840 4850
4801 AACAAATTGTAACTAAGGGAGCCTTTTTCTCAACTAGATAGTATGTACC cohl_al2.seq
4801 AACAAATTGTAACTAAGGGAGCCTTTTTCTCAACTAGATAGTATGTACC a909_al2.seq

TTCTTTCAAGCCTGTAAATGCTAATTATACCACTCTGCTCCTGTTGTATATT Majority
4860 4870 4880 4890 4900
4851 TTCTTTCAAGCCTGTAAATGCTAATTATACCACTCTGCTCCTGTTGTATATT cohl_al2.seq
4851 TTCTTTCAAGCCTGTAAATGCTAATTATACCACTCTGCTCCTGTTGTATATT a909_al2.seq

CTGTTGCCATTAGCTTCTGTGCCCCATTCAACGTTATTTGTTATCGTTAAAG Majority
4910 4920 4930 4940 4950
4901 CTGTTGCCATTAGCTTCTGTGCCCCATTCAACGTTATTTGTTATCGTTAAAG cohl_al2.seq
4901 CTGTTGCCATTAGCTTCTGTGCCCCATTCAACGTTATTTGTTATCGTTAAAG a909_al2.seq

TTTAGAAATTGACCCGTAGCATTCTTTTAAACAAATATAGCACCTTGTA Majority
4960 4970 4980 4990 5000
4951 TTTAGAAATTGACCCGTAGCATTCTTTTAAACAAATATAGCACCTTGTA cohl_al2.seq
4951 TTTAGAAATTGACCCGTAGCATTCTTTTAAACAAATATAGCACCTTGTA a909_al2.seq

TGAAGCTTTTGTGGAACCATCAATTTTTTTTATAGTAATTTGACCATCCC Majority
5010 5020 5030 5040 5050
5001 TGAAGCTTTTGTGGAACCATCAATTTTTTTTATAGTAATTTGACCATCCC cohl_al2.seq
5001 TGAAGCTTTTGTGGAACCATCAATTTTTTTTATAGTAATTTGACCATCCC a909_al2.seq

TCACTGTTACTTTTGAACCTGGCTCATCATTTGCTAGTATTGGGGTTGATG Majority
5060 5070 5080 5090 5100
5051 TCACTGTTACTTTTGAACCTGGCTCATCATTTGCTAGTATTGGGGTTGATG cohl_al2.seq
5051 TCACTGTTACTTTTGAACCTGGCTCATCATTTGCTAGTATTGGGGTTGATG a909_al2.seq

GTGCGCAATGTTTGTATTTTCTGCTAAATCAGCTGAACCTGCTTTAGCTCC Majority
5110 5120 5130 5140 5150
5101 GTGCGCAATGTTTGTATTTTCTGCTAAATCAGCTGAACCTGCTTTAGCTCC cohl_al2.seq
5101 GTGCGCAATGTTTGTATTTTCTGCTAAATCAGCTGAACCTGCTTTAGCTCC a909_al2.seq

ACTCTTTAAATACTCCTGTATAAGTGACTGTGATTGTATTTATTCCCTTAT Majority
5160 5170 5180 5190 5200
5151 ACTCTTTAAATACTCCTGTATAAGTGACTGTGATTGTATTTATTCCCTTAT cohl_al2.seq
5151 ACTCTTTAAATACTCCTGTATAAGTGACTGTGATTGTATTTATTCCCTTAT a909_al2.seq

```

FIGURE 21G

118/487

	AAAAAAAGTCATCATTTAGCTCCCTCCCTTCTGAGTATTTCCGGTTGGAGTATTG Majority	
	5210 5220 5230 5240 5250	
5201	AAAAAAAGTCATCATTTAGCTCCCTTTTGGAGTATTTCCGGTTGGAGTATTG coh1_a12.seq	
5201	AAAAAAAGTCATCATTTAGCTCCCTTTTGGAGTATTTCCGGTTGGAGTATTG a909_a12.seq	
	GTAGCTGCCACGGAATAGTAATCGTGAAATTATTATTTTCTCTAACAAG Majority	
	5260 5270 5280 5290 5300	
5251	GTAGCTGCCACGGAATAGTAATCGTGAAATTATTATTTTCTCTAACAAG coh1_a12.seq	
5251	GTAGCTGCCACGGAATAGTAATCGTGAAATTATTATTTTCTCTAACAAG a909_a12.seq	
	GTTATACTTCCCAGTTGCTTTTTCCGAACCTTGAGTTAGAGTTGTAATAAT Majority	
	5310 5320 5330 5340 5350	
5301	GTTATACTTCCCAGTTGCTTTTTCCGAACCTTGAGTTAGAGTTGTAATAAT coh1_a12.seq	
5301	GTTATACTTCCCAGTTGCTTTTTCCGAACCTTGAGTTAGAGTTGTAATAAT a909_a12.seq	
	TCCCTGATCCATCAGTAATAGTTACTTCATAAGATCCTTCGTTCAAAATCA Majority	
	5360 5370 5380 5390 5400	
5351	TCCCTGATCCATCAGTAATAGTTACTTCATAAGATCCTTCGTTCAAAATCA coh1_a12.seq	
5351	TCCCTGATCCATCAGTAATAGTTACTTCATAAGATCCTTCGTTCAAAATCA a909_a12.seq	
	ACTACAGAAGCAGATGGCATAGTATCCTTTATAACATATTGATACACTTT Majority	
	5410 5420 5430 5440 5450	
5401	ACTACAGAAGCAGATGGCATAGTATCCTTTATAACATATTGATACACTTT coh1_a12.seq	
5401	ACTACAGAAGCAGATGGCATAGTATCCTTTATAACATATTGATACACTTT a909_a12.seq	
	TTCTGTACCATGATAATTGACTGCAATTCCTTATAAGTAATAGTATATTTGA Majority	
	5460 5470 5480 5490 5500	
5451	TTCTGTACCATGATAATTGACTGCAATTCCTTATAAGTAATAGTATATTTGA coh1_a12.seq	
5451	TTCTGTACCATGATAATTGACTGCAATTCCTTATAAGTAATAGTATATTTGA a909_a12.seq	
	CTGTATCACCAACCGAGTACGTTTTTTTGATCTACAGTTTTTCCACCACCA Majority	
	5510 5520 5530 5540 5550	
5501	CTGTATCACCAACCGAGTACGTTTTTTTGATCTACAGTTTTTCCACCACCA coh1_a12.seq	
5501	CTGTATCACCAACCGAGTACGTTTTTTTGATCTACAGTTTTTCCACCACCA a909_a12.seq	
	TCTCCCCATGTCGCATCAGTATTCCTTTTCATCAATAGTAGCATTGGAAGT Majority	
	5560 5570 5580 5590 5600	
5551	TCTCCCCATGTCGCATCAGTATTCCTTTTCATCAATAGTAGCATTGGAAGT coh1_a12.seq	
5551	TCTCCCCATGTCGCATCAGTATTCCTTTTCATCAATAGTAGCATTGGAAGT a909_a12.seq	
	TACAGATGTAACCATAATTACAGCTCCATTATTAACAGTGCTAGAAACAAT Majority	
	5610 5620 5630 5640 5650	
5601	TACAGATGTAACCATAATTACAGCTCCATTATTAACAGTGCTAGAAACAAT coh1_a12.seq	
5601	TACAGATGTAACCATAATTACAGCTCCATTATTAACAGTGCTAGAAACAAT a909_a12.seq	
	AATAATATCCATATTGGGAAACATTAATAACCTCAGTACCATCATTATTT Majority	
	5660 5670 5680 5690 5700	
5651	AATAATATCCATATTGGGAAACATTAATAACCTCAGTACCATCATTATTT coh1_a12.seq	
5651	AATAATATCCATATTGGGAAACATTAATAACCTCAGTACCATCATTATTT a909_a12.seq	
	GACTCAGTAACAGTGGAAACTGCGTGTAGTATTAGCTGATATAGATTTAGC Majority	
	5710 5720 5730 5740 5750	
5701	GACTCAGTAACAGTGGAAACTGCGTGTAGTATTAGCTGATATAGATTTAGC coh1_a12.seq	
5701	GACTCAGTAACAGTGGAAACTGCGTGTAGTATTAGCTGATATAGATTTAGC a909_a12.seq	
	CCATGTCGCAATCTCATTTCCTGACGCAGTATCTTTTTTAGTTACATATG Majority	
	5760 5770 5780 5790 5800	
5751	CCATGTCGCAATCTCATTTCCTGACGCAGTATCTTTTTTAGTTACATATG coh1_a12.seq	
5751	CCATGTCGCAATCTCATTTCCTGACGCAGTATCTTTTTTAGTTACATATG a909_a12.seq	
	TTCTCCCTCCATTAGTAGTTGTCGTAAAAAGAGAAATTAATAATCAGTTGAA Majority	
	5810 5820 5830 5840 5850	
5801	TTCTCCCTCCATTAGTAGTTGTCGTAAAAAGAGAAATTAATAATCAGTTGAA coh1_a12.seq	
5801	TTCTCCCTCCATTAGTAGTTGTCGTAAAAAGAGAAATTAATAATCAGTTGAA a909_a12.seq	

FIGURE 21H

119/487

	G C T T T A T A C T C A G C T T C T T T A C C T T G A G G A A T T A A A T A A G A A G C T C C A T C	Majority
	5860 5870 5880 5890 5900	
5851	G C T T T A T A C T C A G C T T C T T T A C C T T G A G G A A T T A A A T A A G A A G C T C C A T C	cohl_a12.seq
5851	G C T T T A T A C T C A G C T T C T T T A C C T T G A G G A A T T A A A T A A G A A G C T C C A T C	a909_a12.seq
	T T T A T T C G A A T C A G A T A C A T T T G C A T T A T C T A T T T C T G C A T C A A A A A C T T	Majority
	5910 5920 5930 5940 5950	
5901	T T T A T T C G A A T C A G A T A C A T T T G C A T T A T C T A T T T C T G C A T C A A A A A C T T	cohl_a12.seq
5901	T T T A T T C G A A T C A G A T A C A T T T G C A T T A T C T A T T T C T G C A T C A A A A A C T T	a909_a12.seq
	T G T A T G C T T T A T A G G T T G C G C C T T T T T G A G T A T C T T G A A C T G T A A T T G T C	Majority
	5960 5970 5980 5990 6000	
5951	T A T A T G C T T T A T A G G T T G C G C C T T T T T G A G T A T C T T G A A C T G T A A T T G T C	cohl_a12.seq
5951	T G T A T G C T T T A T A G G T T G C G C C T T T T T G A G T A T C T T G A A C T G T A A T T G T C	a909_a12.seq
	C C T G T C T C A G C G G C A A A A G C T A T C G G C G T A A C T G G T G A T A C A G C C A T A C C	Majority
	6010 6020 6030 6040 6050	
6001	C C T G T C T C A G C G G C A A A A G C T A T C G G C G T A A C T G G T G A T A C A G C C A T A C C	cohl_a12.seq
6001	C C T G T C T C A G C G G C A A A A G C T A T C G G C G T A A C T G G T G A T A C A G C C A T A C C	a909_a12.seq
	A A A T G C T A A A C T C G C C A C T A A C A G C G A T T G A A T C A T T T T C T T T T T C A T T G	Majority
	6060 6070 6080 6090 6100	
6051	A A A T G C T A A A C T C G C C A C T A A C A G C G A T T G A A T C A T T T T C T T T T T C A T T G	cohl_a12.seq
6051	A A A T G C T A A A C T C G C C A C T A A C A G C G A T T G A A T C A T T T T C T T T T T C A T T G	a909_a12.seq
	A A A T C T T T T C C T A A A A T C A T A T T G A T G A A T G A T T A A T T C A T A T T T T T T	Majority
	6110 6120 6130 6140 6150	
6101	A A A T C T T T T C C T A A A A T C A T A T T G A T G A A T G A T T A A T T C A T A T T T T T T	cohl_a12.seq
6101	A A A T C T T T T C C T A A A A T C A T A T T G A T G A A T G A T T A A T T C A T A T T T T T T	a909_a12.seq
	T C G A T A G T A T A A T A T T A A T C C T G A T G G T A G A G C T A A A G C T A A A C C A A C T A	Majority
	6160 6170 6180 6190 6200	
6151	T C G A T A G T A T A A T A T T A A T C C T G A T G G T A G A G C T A A A G C T A A A C C A A C T A	cohl_a12.seq
6151	T C G A T A G T A T A A T A T T A A T C C T G A T G G T A G A G C T A A A G C T A A A C C A A C T A	a909_a12.seq
	G C A T A T A A A T G T G T G T T C C A A T A C C T C C A G T A C T A G G C A A T T C T G T T C C T	Majority
	6210 6220 6230 6240 6250	
6201	G C A T A T A A A T G T G T G T T C C A A T A C C T C C A G T A C T A G G C A A T T C T G T T C C T	cohl_a12.seq
6201	G C A T A T A A A T G T G T G T T C C A A T A C C T C C A G T A C T A G G C A A T T C T G T T C C T	a909_a12.seq
	T T A C T G T T A G T A A T T T T A A A A G T A T A T A C T G T A C T T C C A T C T A C T A A A T T	Majority
	6260 6270 6280 6290 6300	
6251	T T A C T G T T A G T A A T T T T A A A A G T A T A T A C T G T A C T T C C A T C T A C T A A A T T	cohl_a12.seq
6251	T T A C T G T T A G T A A T T T T A A A A G T A T A T A C T G T A C T T C C A T C T A C T A A A T T	a909_a12.seq
	C T C T T T T A T T G G T G T C G C A T T A T T A C C A T T T T G T T C A A A G G T A A C T C C C G	Majority
	6310 6320 6330 6340 6350	
6301	C T C T T T T A T T G G T G T C G C A T T A T T A C C A T T T T G T T C A A A G G T A A C T C C C G	cohl_a12.seq
6301	C T C T T T T A T T G G T G T C G C A T T A T T A C C A T T T T G T T C A A A G G T A A C T C C C G	a909_a12.seq
	T A G A A A T C A C T A A T A C T G A T A T A T C A T T T T A G C T A G T A G G T A C C C T G G A	Majority
	6360 6370 6380 6390 6400	
6351	T A G A A A T C A C T A A T A C T G A T A T A T C A T T T T A G C T A G T A G G T A C C C T G G A	cohl_a12.seq
6351	T A G A A A T C A C T A A T A C T G A T A T A T C A T T T T A G C T A G T A G G T A C C C T G G A	a909_a12.seq
	G G G G C C T T T G T C T C T G T T A G G T A G T A T T T T C C T A C T G G C A A A C T G A G G T A	Majority
	6410 6420 6430 6440 6450	
6401	G G G G C C T T T G T C T C T G T T A G G T A G T A T T T T C C T A C T G G C A A A C T G A G G T A	cohl_a12.seq
6401	G G G G C C T T T G T C T C T G T T A G G T A G T A T T T T C C T A C T G G C A A A C T G A G G T A	a909_a12.seq
	G T T A T T A G C A T C C A C T A A T A A C A A G C C T T T A T C G T T T G T C A C C A G C C C T G	Majority
	6460 6470 6480 6490 6500	
6451	G T T A T T A G C A T C C A C T A A T A A C A A G C C T T T A T C G T T T G T C A C C A G C C C T G	cohl_a12.seq
6451	G T T A T T A G C A T C C A C T A A T A A C A A G C C T T T A T C G T T T G T C A C C A G C C C T G	a909_a12.seq

FIGURE 21I

120/487

```

A A T A C A A T T G A T C T G A A G C T T T A T T C C A T T A G C A T C T G A T T C A T A A A T A Majority
6510 6520 6530 6540 6550
6501 A A T A C A T A G G A T G T G A A G C T T T A T T C C C A T T A G C A T C T G A T T C A T A A A T A coh1_a12.seq
6501 A A T A C A T A G G A T G T G A A G C T T T A T T C C C A T T A G C A T C T G A T T C A T A A A T A a909_a12.seq

T C A A A A A A C T G C A C C T G C T A A A A A A T T A T T A T C A T T T T C G A C A T T A A C T T T Majority
6560 6570 6580 6590 6600
6551 T C A A A A A A C T G C A C C T G C T A A A A A A T T A T T A T C A T T T T C G A C A T T A A C T T T coh1_a12.seq
6551 T C A A A A A A C T G C A C C T G C T A A A A A A T T A T T A T C A T T T T C G A C A T T A A C T T T a909_a12.seq

C T G T A G T C G T A C T T T T T G C T T G A T A C G T G T A T T G C T A A A G C T A A T A T C T A Majority
6610 6620 6630 6640 6650
6601 C T G T A G T C G T A C T T T T T G C T T G A T A C G T G T A T T G G T A A A G C T A A T A T C T A coh1_a12.seq
6601 C T G T A G T C G T A C T T T T T G C T T G A T A C G T G T A T T G G T A A A G C T A A T A T C T A a909_a12.seq

C C T C T C C T G A A A C T G T C A G G G A T T G T A A G C C G G T A G C A T C A T A A G T T T T A Majority
6660 6670 6680 6690 6700
6651 C C T C T C C T G A A A C T G T C A G G G A T T G T A A G C C G G T A G C A T C A T A A G T T T T A coh1_a12.seq
6651 C C T C T C C T G A A A C T G T A G G G A T T G T A A G C C G G T A G C A T C A T A A G T T T T A a909_a12.seq

T C A G C T T C A C C A G T T G C T A G A T T T T T T C T G T A A T T G A C T C A G A T A C T T T Majority
6710 6720 6730 6740 6750
6701 T C A G C T T C A C C A G T T G C T A G A T T T T T T C T G T A A T T G A C T C A G A T A C T T T coh1_a12.seq
6701 T C A G C T T C A C C A G T T G C T A G A T T T T T T C T G T A A T T G A C T C A G A T A C T T T a909_a12.seq

A A A T T C A T C G T A G G C T T G T T C A T C T A T T G A T A T A G A A G T T C C A T A A G G T A Majority
6760 6770 6780 6790 6800
6751 A A A T T C A T C G T A G G C T T G T T C A T C T A T T G A T A T A G A A G T T C C A T A A G G T A coh1_a12.seq
6751 A A A T T C A T C G T A G G C T T G T T C A T C T A T T G A T A T A G A A G T T C C A T A A G G T A a909_a12.seq

C T T T A A A T T C C T T A G T C T G A C C A T C T C T C A G C G G A A A A T T C T C T T G T T G C Majority
6810 6820 6830 6840 6850
6801 C T T T A A A T T C C T T A G T C T G A C C A T C T C T C A G C G G A A A A T T C T C T T G T T G C coh1_a12.seq
6801 C T T T A A A T T C C T T A G T C T G A C C A T C T C T C A G C G G A A A A T T C T C T T G T T G C a909_a12.seq

A A C G T T T C A C T T G G A T T A A A C A A G A A G T C T T T C G T C T T A T C T T C A T C T A G Majority
6860 6870 6880 6890 6900
6851 A A C G T T T C A C T T G G A T T A A A C A A G A A G T C T T T C G T C T T A T C T T C A T C T A G coh1_a12.seq
6851 A A C G T T T C A C T T G G A T T A A A C A A G A A G T C T T T C G T C T T A T C T T C A T C T A G a909_a12.seq

T C C A A C G A C A G T T T T A C T T A C T C T G A C G G T G T A T T C T T T A G G T T G C C A A A Majority
6910 6920 6930 6940 6950
6901 T C C A A C G A C A G T T T T A C T T A C T C T G A C G G T G T A T T C T T T A G G T T G C C A A A coh1_a12.seq
6901 T C C A A C G A C A G T T T T A C T T A C T C T G A C G G T G T A T T C T T T A G G T T G C C A A A a909_a12.seq

C A G C A T A T A A G G T A T T T G T T G C A T C A G G G T T G T T A T C A A T A C C T A T T G A T Majority
6960 6970 6980 6990 7000
6951 C A G C A T A T A A G G T A T T T G T T G C A T C A G G G T T G T T A T C A A T A C C T A T T G A T coh1_a12.seq
6951 C A G C A T A T A A G G T A T T T G T T G C A T C A G G G T T G T T A T C A A T A C C T A T T G A T a909_a12.seq

T G A C C T G C T G T A A A T T C C A C A C G T C C T G T A T C A G C T A A A T C C T T A T C A T G Majority
7010 7020 7030 7040 7050
7001 T G A C C T G C T G T A A A T T C C A C A C G T C C T G T A T C A G C T A A A T C C T T A T C A T G coh1_a12.seq
7001 T G A C C T G C T G T A A A T T C C A C A C G T C C T G T A T C A G C T A A A T C C T T A T C A T G a909_a12.seq

A T G C C A A C C A A T A A G G T T G T A A C C T G T C C T T G T A A A G T A T T G G T T T T C A G Majority
7060 7070 7080 7090 7100
7051 A T G C C A A C C A A T A A G G T T G T A A C C T G T C C T T G T A A A G T A T T G G T T T T C A G coh1_a12.seq
7051 A T G C C A A C C A A T A A G G T T G T A A C C T G T C C T T G T A A A G T A T T G G T T T T C A G a909_a12.seq

G A A T T G T A G T T G T G C T A T T C A A C T C C A T A C G C G G T G T C T C T A C T T G T G T T Majority
7110 7120 7130 7140 7150
7101 G A A T T G T A G T T G T G C T A T T C A A C T C C A T A C G C G G T G T C T C T A C T T G T G T T coh1_a12.seq
7101 G A A T T G T A G T T G T G C T A T T C A A C T C C A T A C G C G G T G T C T C T A C T T G T G T T a909_a12.seq

```

FIGURE 21J

121/487

	ACCACGATACCACTTCTAGTACCACCGTTACCAATTGTAATTTGAT	Majority
	7160 7170 7180 7190 7200	
7151	ACCACATTACCACTTCTAGTACCACCGTTACCAATTGTAATTTGAT	cohl_al2.seq
7151	ACCACATTACCACTTCTAGTACCACCGTTACCAATTGTAATTTGAT	a909_al2.seq
	TCAGGTAATCTTCTAATTTGATATCTCCTACTGGAATAATGACAGGTTTTA	Majority
	7210 7220 7230 7240 7250	
7201	TCAGGTAATCTTCTAATTTGATATCTCCTACTGGAATAATGACAGGTTTTA	cohl_al2.seq
7201	TCAGGTAATCTTCTAATTTGATATCTCCTACTGGAATAATGACAGGTTTTA	a909_al2.seq
	TGGTGATATTTTTATTAGCATCTGCTAAATGGGCGTCAATATCAATGGAA	Majority
	7260 7270 7280 7290 7300	
7251	TGGTGATATTTTTATTAGCATCTGCTAAATGGGCGTCAATATCAATGGAA	cohl_al2.seq
7251	TGGTGATATTTTTATTAGCATCTGCTAAATGGGCGTCAATATCAATGGAA	a909_al2.seq
	TCATATGGGTTATAAATTTTACCATTGTACCACCAGCCACGGAAACGATA	Majority
	7310 7320 7330 7340 7350	
7301	TCATATGGGTTATAAATTTTACCATTGTACCACCAGCCACGGAAACGATA	cohl_al2.seq
7301	TCATATGGGTTATAAATTTTACCATTGTACCACCAGCCACGGAAACGATA	a909_al2.seq
	GCCATCTGGCATTGTCGGACGTCTCAGTAAGGCTGAATGGGAGCCATCGT	Majority
	7360 7370 7380 7390 7400	
7351	GCCATCTGGCATTGTCGGACGTCTCAGTAAGGCTGAATGGGAGCCATCGT	cohl_al2.seq
7351	GCCATCTGGCATTGTCGGACGTCTCAGTAAGGCTGAATGGGAGCCATCGT	a909_al2.seq
	CATAGGAATCAGGGTCAGTAGGCTCATTACTTGTGTAATTGCTGACCA	Majority
	7410 7420 7430 7440 7450	
7401	CATAGGAATCAGGGTCAGTAGGCTCATTACTTGTGTAATTGCTGACCA	cohl_al2.seq
7401	CATAGGAATCAGGGTCAGTAGGCTCATTACTTGTGTAATTGCTGACCA	a909_al2.seq
	GAAGCATCCAATGCTGGCTTTCCATCTGTACCAACAGCATCATTGCTGTA	Majority
	7460 7470 7480 7490 7500	
7451	GAAGCATCCAATGCTGGCTTTCCATCTGTACCAACAGCATCATTGCTGTA	cohl_al2.seq
7451	GAAGCATCCAATGCTGGCTTTCCATCTGTACCAACAGCATCATTGCTGTA	a909_al2.seq
	TATAATATGATAAATCTCCAGCCCTTTGCCCAAATAGCTCTTAAATTGATAT	Majority
	7510 7520 7530 7540 7550	
7501	TATAATATGATAAATCTCCAGCCCTTTGCCCAAATAGCTCTTAAATTGATAT	cohl_al2.seq
7501	TATAATATGATAAATCTCCAGCCCTTTGCCCAAATAGCTCTTAAATTGATAT	a909_al2.seq
	CTTGAGTTACAGCACCTGAAAAGTTATAAGGTCTAATACTACCATCTGGA	Majority
	7560 7570 7580 7590 7600	
7551	CTTGAGTTACAGCACCTGAAAAGTTATAAGGTCTAATACTACCATCTGGA	cohl_al2.seq
7551	CTTGAGTTACAGCACCTGAAAAGTTATAAGGTCTAATACTACCATCTGGA	a909_al2.seq
	TTAACATAAATACCAACCGACTAATTTGTAAAGCGTCTTTTACGTACTTGTA	Majority
	7610 7620 7630 7640 7650	
7601	TTAACATAAATACCAACCGACTAATTTGTAAAGCGTCTTTTACGTACTTGTA	cohl_al2.seq
7601	TTAACATAAATACCAACCGACTAATTTGTAAAGCGTCTTTTACGTACTTGTA	a909_al2.seq
	CTTAGTTGTTGTAATCAACATTTGAGAGACTAGTATCTGTCGTATAAATAGG	Majority
	7660 7670 7680 7690 7700	
7651	CTTAGTTGTTGTAATCAACATTTGAGAGACTAGTATCTGTCGTATAAATAGG	cohl_al2.seq
7651	CTTAGTTGTTGTAATCAACATTTGAGAGACTAGTATCTGTCGTATAAATAGG	a909_al2.seq
	CATCTTTAGTTGAGTGGGATCTTTATCTCGTGAATCATACTTATAATAA	Majority
	7710 7720 7730 7740 7750	
7701	CATCTTTAGTTGAGTGGGATCTTTATCTCGTGAATCATACTTATAATAA	cohl_al2.seq
7701	CATCTTTAGTTGAGTGGGATCTTTATCTCGTGAATCATACTTATAATAA	a909_al2.seq
	TATGTACCTGAAGCATCTTGGATATAAATCCCTTGTAATATCTGTATAATC	Majority
	7760 7770 7780 7790 7800	
7751	TATGTACCTGAAGCATCTTGGATATAAATCCCTTGTAATATCTGTATAATC	cohl_al2.seq
7751	TATGTACCTGAAGCATCTTGGATATAAATCCCTTGTAATATCTGTATAATC	a909_al2.seq

FIGURE 21K

122/487

```

CGGAAATACGATCACCATAATGCAAATCTAAATAGGTATCATCTGTTTTTC Majority
7801 CGGAAATACGATCACCATAATGCAAATCTAAATAGGTATCATCTGTTTTTC cohl_al2.seq
7801 CGGAAATACGATCACCATAATGCAAAGTCTAAATAGGTATCATCTGTTTTTC a909_al2.seq

ATAATTGGCCTCCGTTTGGATCAATATTGACACCATATGTTACCTTTTGC Majority
7860 ATAATTGGCCTCCGTTTGGATCAATATTGACACCATATGTTACCTTTTGC cohl_al2.seq
7851 ATAATTGGCCTCCGTTTGGATCAATATTGACACCATATGTTACCTTTTGC a909_al2.seq

CAACCTGCGATAGACTTTTAACATCATGAGGAGGCATAGTCCGTGTTAAAGTGC Majority
7910 CAACCTGCGATAGACTTTTAACATCATGAGGAGGCATAGTCCGTGTTAAAGTGC cohl_al2.seq
7901 CAACCTGCGATAGACTTTTAACATCATGAGGAGGCATAGTCCGTGTTAAAGTGC a909_al2.seq

AAATACTTGTGTTTTGTGCTTGGTCTTTTATACCATTTACCATCCCAAACAT Majority
7960 AAATACTTGTGTTTTGTGCTTGGTCTTTTATACCATTTACCATCCCAAACAT cohl_al2.seq
7951 AAATACTTGTGTTTTGTGCTTGGTCTTTTATACCATTTACCATCCCAAACAT a909_al2.seq

ACCTGCTCGACTAGGTTTAGGTTGAACCGTTGTCGTATCGGGGGGCATAA Majority
8010 ACCTGCTCGACTAGGTTTAGGTTGAACCGTTGTCGTATCGGGGGGCATAA cohl_al2.seq
8001 ACCTGCTCGACTAGGTTTAGGTTGAACCGTTGTCGTATCGGGGGGCATAA a909_al2.seq

GAGGACAAATTTTGTCTCATATAGAACATCCCTTACTGGAAAATTAGGAAG Majority
8060 GAGGACAAATTTTGTCTCATATAGAACATCCCTTACTGGAAAATTAGGAAG cohl_al2.seq
8051 GAGGACAAATTTTGTCTCATATAGAACATCCCTTACTGGAAAATTAGGAAG a909_al2.seq

CTCTGTATTATCAAGCGGATCTAAATATTTAATCTTGTATGAATTACGTT Majority
8110 CTCTGTATTATCAAGCGGATCTAAATATTTAATCTTGTATGAATTACGTT cohl_al2.seq
8101 CTCTGTATTATCAAGCGGATCTAAATATTTAATCTTGTATGAATTACGTT a909_al2.seq

CATACCATAACCACTAAGTTCAAATAAATCTTTGCTGCTAGTCTCCATATTTA Majority
8160 CATACCATAACCACTAAGTTCAAATAAATCTTTGCTGCTAGTCTCCATATTTA cohl_al2.seq
8151 CATACCATAACCACTAAGTTCAAATAAATCTTTGCTGCTAGTCTCCATATTTA a909_al2.seq

TCGTAGTATTTCATCTCGGATTGGCACCTTTTGTTTTTGCACTCTGTTTGTCT Majority
8210 TCGTAGTATTTCATCTCGGATTGGCACCTTTTGTTTTTGCACTCTGTTTGTCT cohl_al2.seq
8201 TCGTAGTATTTCATCTCGGATTGGCACCTTTTGTTTTTGCACTCTGTTTGTCT a909_al2.seq

TGGGTTCTGATCAAATAAGGTAATTATCTGGATATAAGCTTTTGATAGTATT Majority
8260 TGGGTTCTGATCAAATAAGGTAATTATCTGGATATAAGCTTTTGATAGTATT cohl_al2.seq
8251 TGGGTTCTGATCAAATAAGGTAATTATCTGGATATAAGCTTTTGATAGTATT a909_al2.seq

TAACATTAAATCCTAGGTATTTTTCTGTAAAGGTAATTCCTCTGCTCCA Majority
8310 TAACATTAAATCCTAGGTATTTTTCTGTAAAGGTAATTCCTCTGCTCCA cohl_al2.seq
8301 TAACATTAAATCCTAGGTATTTTTCTGTAAAGGTAATTCCTCTGCTCCA a909_al2.seq

GCACCTCCCTCTGTCTGCTAAAGACTATTTGCCATCTAGTCTTCTTTT Majority
8360 GCACCTCCCTCTGTCTGCTAAAGACTATTTGCCATCTAGTCTTCTTTT cohl_al2.seq
8351 GCACCTCCCTCTGTCTGCTAAAGACTATTTGCCATCTAGTCTTCTTTT a909_al2.seq

GTAGAACGGATAATTTTGAATTCTCTTCCCTTTTGGATAGAGTTTTATT Majority
8410 GTAGAACGGATAATTTTGAATTCTCTTCCCTTTTGGATAGAGTTTTATT cohl_al2.seq
8401 GTAGAACGGATAATTTTGAATTCTCTTCCCTTTTGGATAGAGTTTTATT a909_al2.seq

```

FIGURE 21L



124/487

	G T C C A A G C A G G T T G G T C T G T T A A C A C A C G A C G A G G T G C A A T A T A G G T A A C Majority									
	9110	9120	9130	9140	9150					
9089	G T C C A A G C A G G T T G G T C T G T T A A C A C A C G A C G A G G T G C A A T A T A G G T A A C coh1_al2.seq									
9086	G T C C A A G C A G G T T G G T C T G T T A A C A C A C G A C G A G G T G C A A T A T A G G T A A C a909_al2.seq									
	T C C T G T T G A T A A G T C G C C T G T G T T G A A T T C A A C A C C G T A C G A T T C T T T A A Majority									
	9160	9170	9180	9190	9200					
9139	T C C T G T T G A T A A G T C G C C T G T G T T G A A T T C A A C A C C G T A C G A T T C T T T A A coh1_al2.seq									
9136	T C C T G T T G A T A A G T C G C C T G T G T T G A A T T C A A C A C C G T A C G A T T C T T T A A a909_al2.seq									
	A G A C A G G A T A A A G A C T T A T G T C T G A A G A T A C A G G T A A T G C T T G A A T T T C T Majority									
	9210	9220	9230	9240	9250					
9189	A G A C A G G A T A A A G A C T T A T G T C T G A A G A T A C A G G T A A T G C T T G A A T T T C T coh1_al2.seq									
9186	A G A C A G G A T A A A G A C T T A T G T C T G A A G A T A C A G G T A A T G C T T G A A T T T C T a909_al2.seq									
	G A T T C A G A A A G T G G A G C T C C A T T T T G A G T T T T A G A C C A A C C T A C A A A T A A Majority									
	9260	9270	9280	9290	9300					
9239	G A T T C A G A A A G T G G A G C T C C A T T T T G A G T T T T A G A C C A A C C T A C A A A T A A coh1_al2.seq									
9236	G A T T C A G A A A G T G G A G C T C C A T T T T G A G T T T T A G A C C A A C C T A C A A A T A A a909_al2.seq									
	T A A T G T T G A A G T G G G T G G T G A A A C T T T A A A T G A A C T T A G A T C C G C G G T T C Majority									
	9310	9320	9330	9340	9350					
9289	T A A T G T T G A A G T G G G T G G T G A A A C T T T A A A T G A A C T T A A T C C G C G G T T C coh1_al2.seq									
9286	T A A T G T T G A A G T G G G T G G T G A A A C T T T A A A T G A A C T T A G A T C C G C G G T T C a909_al2.seq									
	C G T T T C C T G A A C T A T C C A A T G G T A C T T G A T A A C G C T C C A A A A T A C T T T T A Majority									
	9360	9370	9380	9390	9400					
9339	C G T T T C C T G A A C T A T C C A A T G G T A C T T G A T A A C G C T C C A A A A T A C T T T T A coh1_al2.seq									
9336	C G T T T C C T G A A C T A T C C A A T G G T A C T T G A T A A C G C T C C A A A A T A C T T T T A a909_al2.seq									
	C C A G C A G A A T C A T C A T A T A A A G T T A C T G T C G C C A C T T T A C T C A T A A A T G C Majority									
	9410	9420	9430	9440	9450					
9389	C C A G C A G A A T C A T C A T A T A A A G T T A C T G T C G C C A C T T T A C T C A T A A A T G C coh1_al2.seq									
9386	C C A G C A G A A T C A T C A T A T A A A G T T A C T G T C G C C A C T T T A C T C A T A A A T G C a909_al2.seq									
	A C G A A C A T A A A T T T C T T T T G T C T C A G T T A C A G T T A T T G G C T C A C C A A A T T Majority									
	9460	9470	9480	9490	9500					
9439	A C G A A C A T A A A T T T C T T T T G T C T C A G T T A C A G T T A T T G G C T C A C C A A A T T coh1_al2.seq									
9436	A C G A A C A T A A A T T T C T T T T G T C T C A G T T A C A G T T A T T G G C T C A C C A A A T T a909_al2.seq									
	T A A C A G G G T C A C C A T A C T T T C C A G T A G T A G G A T C A T A G G T A T A C C A A C C A Majority									
	9510	9520	9530	9540	9550					
9489	T A A C A G G G T C A C C A T A C T T T C C A G T A G T A G G A T C A T A G G T A T A C C A A C C A coh1_al2.seq									
9486	T A A C A G G G T C A C C A T A C T T T C C A G T A G T A G G A T C A T A G G T A T A C C A A C C A a909_al2.seq									
	T T A A A A T G C T C T C C T G C T T T A A T C G T C G G A A T C C C A A C T T C T C C T A G A G A Majority									
	9560	9570	9580	9590	9600					
9539	T T A A A A T G C T C T C C T G C T T T A A T C G T C G G A A T C C C A A C T T C T C C T A G A G A coh									

FIGURE 21N

FIGURE 210

FIGURE 21P

127/487

CGCCCTCTCTTATCTACTAAATTTCTTTACGTTTTCTAAATAGGTATCTGCA Majority
11060 11070 11080 11090 11100
11039 CGCCCTCTCTTATCTACTAAATTTCTTTACGTTTTCTAAATAGGTATCTGCA coh1_a12.seq
11035 CGCCCTCTCTTATCTACTAAATTTCTTTACGTTTTCTAAATAGGTATCTGCA a909_a12.seq

GCTAACTGAGCAGTCTCAAAAAATACCACTAAGCATTAAAGGAAGCTTCGGC Majority
11110 11120 11130 11140 11150
11089 GCTAACTGAGCAGTCTCAAAAAATACCACTAAGCATTAAAGGAAGCTTCGGC coh1_a12.seq
11085 GCTAACTGAGCAGTCTCAAAAAATACCACTAAGCATTAAAGGAAGCTTCGGC a909_a12.seq

AATAGAACCAGCTTTTGCTAATTTTGAATTTTTTATCATCTAGCGCTTCTT Majority
11160 11170 11180 11190 11200
11139 AATAGAACCAGCTTTTGCTAATTTTGAATTTTTTATCATCTAGCGCTTCTT coh1_a12.seq
11135 AATAGAACCAGCTTTTGCTAATTTTGAATTTTTTATCATCTAGCGCTTCTT a909_a12.seq

TAACTTGCTGAATCTCTTTCTCTTGTCTTTTCAATAAGAAGTTGCTGTTCT Majority
11210 11220 11230 11240 11250
11189 TAACTTGCTGAATCTCTTTCTCTTGTCTTTTCAATAAGAAGTTGCTGTTCT coh1_a12.seq
11185 TAACTTGCTGAATCTCTTTCTCTTGTCTTTTCAATAAGAAGTTGCTGTTCT a909_a12.seq

AACATAAATTTCTAGCAAGTCTTTTCTTTTAAATTTTTTAAATCTTCCAT Majority
11260 11270 11280 11290 11300
11239 AACATAAATTTCTAGCAAGTCTTTTCTTTTAAATTTTTTAAATCTTCCAT coh1_a12.seq
11235 AACATAAATTTCTAGCAAGTCTTTTCTTTTAAATTTTTTAAATCTTCCAT a909_a12.seq

CGCGATTACTTCTTAACTGAACCTTAAATTATCGTTTAGATATTATATC Majority
11310 11320 11330 11340 11350
11289 CGCGATTACTTCTTAACTGAACCTTAAATTATCGTTTAGATATTATATC coh1_a12.seq
11285 CGCGATTACTTCTTAACTGAACCTTAAATTATCGTTTAGATATTATATC a909_a12.seq

AAAGTTCTAACCCTTTAAACTCATTTTTTGTCTGCTTTTTTCTCAAAAAA Majority
11360 11370 11380 11390 11400
11339 AAAGTTCTAACCCTTTAAACTCATTTTTTGTCTGCTTTTTTCTCAAAAAA coh1_a12.seq
11335 AAAGTTCTAACCCTTTAAACTCATTTTTTGTCTGCTTTTTTCTCAAAAAA a909_a12.seq

AGTCTATGCTAAATTAACATTTTTGATAATTTTTTGA AAAATCTCATCGA Majority
11410 11420 11430 11440 11450
11389 AGTCTATGCTAAATTAACATTTTTGATAATTTTTTGA AAAATCTCATCGA coh1_a12.seq
11385 AGTCTATGCTAAATTAACATTTTTGATAATTTTTTGA AAAATCTCATCGA a909_a12.seq

AGTCATTTTCTTTTTTGAAGCTCGAATTCTAGGCATTAAAAAGCCATATA Majority
11460 11470 11480 11490 11500
11439 AGTCATTTTCTTTTTTGAAGCTCGAATTCTAGGCATTAAAAAGCCATATA coh1_a12.seq
11435 AGTCATTTTCTTTTTTGAAGCTCGAATTCTAGGCATTAAAAAGCCATATA a909_a12.seq

TC AAAATTGATATATGGCTTTTTTTTATTATTTAAAAACAAAAGCAATCAATAG Majority
11510 11520 11530 11540 11550
11489 TC AAAATTGATATATGGCTTTTTTTTATTATTTAAAAACAAAAGCAATCAATAG coh1_a12.seq
11476 TC AAAATTGATATATGGCTTTTTTTTATTATTTAAAAACAAAAGCAATCAATAG a909_a12.seq

GACAATAGCGTCAATTTAGTGACATAATCTATTACAGATTAAAGTTCTTTT Majority
11560 11570 11580 11590 11600
11539 GACAATAGCGTCAATTTAGTGACATAATCTATTACAGATTAAAGTTCTTTT coh1_a12.seq
11492 GACAATAGCGTCAATTTAGTGACATAATCTATTACAGATTAAAGTTCTTTT a909_a12.seq

TGAATAATATAATCCAACTTTTCAACTGTTTTTTCCCATGTGAAATGTTT Majority
11610 11620 11630 11640 11650
11589 TGAATAATATAATCCAACTTTTCAACTGTTTTTTCCCATGTGAAATGTTT coh1_a12.seq
11542 TGAATAATATAATCCAACTTTTCAACTGTTTTTTCCCATGTGAAATGTTT a909_a12.seq

TTTAATTCTTTTAGCAATATTCTGTTGTAGTTTCTCTCTTAATGCCTTAT Majority
11660 11670 11680 11690 11700
11639 TTTAATTCTTTTAGCAATATTCTGTTGTAGTTTCTCTCTTAATGCCTTAT coh1_a12.seq
11592 TTTAATTCTTTTAGCAATATTCTGTTGTAGTTTCTCTCTTAATGCCTTAT a909_a12.seq

FIGURE 21Q

128/487

```

CTTTTAACTAAATAAATAAGAGATTCTCTGAGTGAAGTATTTTCTTCC Majority
11710 11720 11730 11740 11750
11689 CTTTACTAATAAATCAAGAGATTTCATGGAGTGAAGTATTTTCTTCC cohl_a12.seq
11642 CTTTACTAATAAATCAAGAGATTTCATGGAGTGAAGTATTTTCTTCC a909_a12.seq

ATGATGATTCTTAAGTCAAGGCTATCAATAAATTCAACTGTTCCACCGCC Majority
11760 11770 11780 11790 11800
11739 ATGATGATTCTTAAGTCAAGGCTATCAATAAATTCAACTGTTCCACCGCC cohl_a12.seq
11692 ATGATGATTCTTAAGTCAAGGCTATCAATAAATTCAACTGTTCCACCGCC a909_a12.seq

ATCTGTTGCAATAATAGCACTCGAAAGTAGACCAGCTTCTAAAAATAGAGG Majority
11810 11820 11830 11840 11850
11789 ATCTGTTGCAATAATAGCACTCGAAAGTAGACCAGCTTCTAAAAATAGAGG cohl_a12.seq
11742 ATCTGTTGCAATAATAGCACTCGAAAGTAGACCAGCTTCTAAAAATAGAGG a909_a12.seq

TTGGTAATCCCTCTCGGATACATTGATGGGTAAACAAAGATATCTGTCTGT Majority
11860 11870 11880 11890 11900
11839 TTGGTAATCCCTCTCGGATACATTGATGGGTAAACAAAGATATCTGTCTGT cohl_a12.seq
11792 TTGGTAATCCCTCTCGGATACATTGATGGGTAAACAAAGATATCTGTCTGT a909_a12.seq

GCCATTAAAGACATAGTCTGTTCAAAGTTTAAATTTCCCAAAAAAGTTAAT Majority
11910 11920 11930 11940 11950
11889 GCCATTAAAGACATAGTCTGTTCAAAGTTTAAATTTCCCAAAAAAGTTAAT cohl_a12.seq
11842 GCCATTAAAGACATAGTCTGTTCAAAGTTTAAATTTCCCAAAAAAGTTAAT a909_a12.seq

CTGTTTGGAAGTATATTTCTCTTTCAAATGTGCTAATTCAGGTCCGTCTC Majority
11960 11970 11980 11990 12000
11939 CTGTTTGGAAGTATATTTCTCTTTCAAATGTGCTAATTCAGGTCCGTCTC cohl_a12.seq
11892 CTGTTTGGAAGTATATTTCTCTTTCAAATGTGCTAATTCAGGTCCGTCTC a909_a12.seq

CTGCAATCTGTAAATAAACATTTTTCAGAGTACTGTGACATCGAAAAATGCT Majority
12010 12020 12030 12040 12050
11989 CTGCAATCTGTAAATAAACATTTTTCAGAGTACTGTGACATCGAAAAATGCT cohl_a12.seq
11942 CTGCAATCTGTAAATAAACATTTTTCAGAGTACTGTGACATCGAAAAATGCT a909_a12.seq

TCTAAGAGCAATTCAATGCCCTTTTCTTTAATAATTCTACCAGCATAAGT Majority
12060 12070 12080 12090 12100
12039 TCTAAGAGCAATTCAATGCCCTTTTCTTTAATAATTCTACCAGCATAAGT cohl_a12.seq
11992 TCTAAGAGCAATTCAATGCCCTTTTCTTTAATAATTCTACCAGCATAAGT a909_a12.seq

GATGAAAAATATCATCAGCAGATTTTTCAGGTAAGCCGTGTACGCAAAAAT Majority
12110 12120 12130 12140 12150
12089 GATGAAAAATATCATCAGCAGATTTTTCAGGTAAGCCGTGTACGCAAAAAT cohl_a12.seq
12042 GATGAAAAATATCATCAGCAGATTTTTCAGGTAAGCCGTGTACGCAAAAAT a909_a12.seq

CAGAGCCTAGACTTTTCAGATACCGAATTATAAATAACTCTTTAGCTTCT Majority
12160 12170 12180 12190 12200
12139 CAGAGCCTAGACTTTTCAGATACCGAATTATAAATAACTCTTTAGCTTCT cohl_a12.seq
12092 CAGAGCCTAGACTTTTCAGATACCGAATTATAAATAACTCTTTAGCTTCT a909_a12.seq

ATATTTAAATGTTTTAACCATTCAACGGCTTCTCTTGGATACCGCATAAAA Majority
12210 12220 12230 12240 12250
12189 ATATTTAAATGTTTTAACCATTCAACGGCTTCTCTTGGATACCGCATAAAA cohl_a12.seq
12142 ATATTTAAATGTTTTAACCATTCAACGGCTTCTCTTGGATACCGCATAAAA a909_a12.seq

ATCTGGACCATAGTCTTAACACGGCGCTGTGAGAGATGTTTCATAGATAG Majority
12260 12270 12280 12290 12300
12239 ATCTGGACCATAGTCTTAACACGGCGCTGTGAGAGATGTTTCATAGATAG cohl_a12.seq
12192 ATCTGGACCATAGTCTTAACACGGCGCTGTGAGAGATGTTTCATAGATAG a909_a12.seq

CTCCAAAGAAATCTAAAAAACGATTATTGACAGAAAAATGACTTGACCCA Majority
12310 12320 12330 12340 12350
12289 CTCCAAAGAAATCTAAAAAACGATTATTGACAGAAAAATGACTTGACCCA cohl_a12.seq
12242 CTCCAAAGAAATCTAAAAAACGATTATTGACAGAAAAATGACTTGACCCA a909_a12.seq

```

FIGURE 21R

129/487

	<u>TGGTCTAAAAACAATACTAGGTAACGGTAAACGGTGGCTTCTTTGCAAAAGATAGCCC</u>	Majority
	12360 12370 12380 12390 12400	
12339	TGGTCTAAAAACAATACTAGGTAACGGTGGCTTCTTTGCAAAAGATAGCCC	coh1_a12.seq
12292	TGGTCTAAAAACAATACTAGGTAACGGTGGCTTCTTTGCAAAAGATAGCCC	a909_a12.seq
	<u>TTCTAGCGTTGTTAACTGAAAACGCTGATTACAAATCACAAAATCAATAT</u>	Majority
	12410 12420 12430 12440 12450	
12389	TTCTAGCGTTGTTAACTGAAAACGCTGATTACAAATCACAAAATCAATAT	coh1_a12.seq
12342	TTCTAGCGTTGTTAACTGAAAACGCTGATTACAAATCACAAAATCAATAT	a909_a12.seq
	<u>TTTCATCTGAAAACATGTTTCATCAGCGTGTGTTGTTCTCGATTTTTTGTTA</u>	Majority
	12460 12470 12480 12490 12500	
12439	TTTCATCTGAAAACATGTTTCATCAGCGTGTGTTGTTCTCGATTTTTTGTTA	coh1_a12.seq
12392	TTTCATCTGAAAACATGTTTCATCAGCGTGTGTTGTTCTCGATTTTTTGTTA	a909_a12.seq
	<u>ATAATAGGATAGCGCTGCTTGACAAAGTTTTTGGTCGGTAAACCGGTAAAT</u>	Majority
	12510 12520 12530 12540 12550	
12489	ATAATAGGATAGCGCTGCTTGACAAAGTTTTTGGTCGGTAAACCGGTAAAT	coh1_a12.seq
12442	ATAATAGGATAGCGCTGCTTGACAAAGTTTTTGGTCGGTAAACCGGTAAAT	a909_a12.seq
	<u>TTTTCTACCCTTGCTTTCATCTATAATCGGTAAATCATCATGATTAGTTG</u>	Majority
	12560 12570 12580 12590 12600	
12539	TTTTCTACCCTTGCTTTCATCTATAATCGGTAAATCATCATGATTAGTTG	coh1_a12.seq
12492	TTTTCTACCCTTGCTTTCATCTATAATCGGTAAATCATCATGATTAGTTG	a909_a12.seq
	<u>TTACAATAACAACACGGTAGCCACGCTTAACCAATCTGCTGTCATTTTA</u>	Majority
	12610 12620 12630 12640 12650	
12589	TTACAATAACAACACGGTAGCCACGCTTAACCAATCTGCTGTCATTTTA	coh1_a12.seq
12542	TTACAATAACAACACGGTAGCCACGCTTAACCAATCTGCTGTCATTTTA	a909_a12.seq
	<u>TCTGTATAACGTTCAATACCTCCGAGGAAGGGTAGATAATATCCTGAGAA</u>	Majority
	12660 12670 12680 12690 12700	
12639	TCTGTATAACGTTCAATACCTCCGAGGAAGGGTAGATAATATCCTGAGAA	coh1_a12.seq
12592	TCTGTATAACGTTCAATACCTCCGAGGAAGGGTAGATAATATCCTGAGAA	a909_a12.seq
	<u>AACAGCAACTGTTTTTACCTTATTTTCCATATTTATCCACTTTTCATCAAT</u>	Majority
	12710 12720 12730 12740 12750	
12689	AACAGCAACTGTTTTTACCTTATTTTCCATATTTATCCACTTTTCATCAAT	coh1_a12.seq
12642	AACAGCAACTGTTTTTACCTTATTTTCCATATTTATCCACTTTTCATCAAT	a909_a12.seq
	<u>AAGCCATCTTTTTAAGCCTTTAATCATAGCAACTAATTTTTTGGCTCTTTTG</u>	Majority
	12760 12770 12780 12790 12800	
12739	AAGCCATCTTTTTAAGCCTTTAATCATAGCAACTAATTTTTTGGCTCTTTTG	coh1_a12.seq
12692	AAGCCATCTTTTTAAGCCTTTAATCATAGCAACTAATTTTTTGGCTCTTTTG	a909_a12.seq
	<u>CTCTTCTGCTACCAACACTCGAACAATTCATTTEGCATAAAATACTAAAT</u>	Majority
	12810 12820 12830 12840 12850	
12789	CTCTTCTGCTACCAACACTCGAACAATTCATTTEGCATAAAATACTAAAT	coh1_a12.seq
12742	CTCTTCTGCTACCAACACTCGAACAATTCATTTEGCATAAAATACTAAAT	a909_a12.seq
	<u>ATTGTGCGCGTTCTTCTTACCATATTTTATATAATATAAATCGCATTC</u>	Majority
	12860 12870 12880 12890 12900	
12839	ATTGTGCGCGTTCTTCTTACCATATTTTATATAATATAAATCGCATTC	coh1_a12.seq
12792	ATTGTGCGCGTTCTTCTTACCATATTTTATATAATATAAATCGCATTC	a909_a12.seq
	<u>CGTATCATGTAATATTTTTCGAAATGGTGAATGATTCAATACATGAAAAAC</u>	Majority
	12910 12920 12930 12940 12950	
12889	CGTATCATGTAATATTTTTCGAAATGGTGAATGATTCAATACATGAAAAAC	coh1_a12.seq
12842	CGTATCATGTAATATTTTTCGAAATGGTGAATGATTCAATACATGAAAAAC	a909_a12.seq
	<u>ATGCCCAAATTTTTTAACTCGTGAAGAGTGTCGAATTTTCGTGTAAAAGAC</u>	Majority
	12960 12970 12980 12990 13000	
12939	ATGCCCAAATTTTTTAACTCGTGAAGAGTGTCGAATTTTCGTGTAAAAGAC	coh1_a12.seq
12892	ATGCCCAAATTTTTTAACTCGTGAAGAGTGTCGAATTTTCGTGTAAAAGAC	a909_a12.seq

FIGURE 21S

130/487

	C A A T A T A A T T A A C C T G A T A A G T C T T A T A T C C C A T C T C T G A C A G A C G A T A A	Majority
	13010 13020 13030 13040 13050	
12989	C A A T A A A A T T A A C C T G A T A A G T C T T A T A T C C C A T C T C T G A C A G A C G A T A A	cohl_al2.seq
12942	C A A T A A A A T T A A C C T G A T A A G T C T T A T A T C C C A T C T C T G A C A G A C G A T A A	a909_al2.seq
	T T C A T T T C A G A G T C A A C A A A A T C A A T A A A C A T C T C T T C T G C A A A G C C A G A	Majority
	13060 13070 13080 13090 13100	
13039	T T C A T T T C A G A G T C A A C A A A A T C A A T A A A C A T C T C T T C T G C A A A G C C A G A	cohl_al2.seq
12992	T T C A T T T C A G A G T C A A C A A A A T C A A T A A A C A T C T C T T C T G C A A A G C C A G A	a909_al2.seq
	T G T T T C T T C A A A A A C G C T C G T T T T C A T T A A A G C A G C C G A A G T A A T A C A C T	Majority
	13110 13120 13130 13140 13150	
13089	T G T T T C T T C A A A A A C G C T C G T T T T C A T T A A A G C A G C C G A A G T A A T A C A C T	cohl_al2.seq
13042	T G T T T C T T C A A A A A C G C T C G T T T T C A T T A A A G C A G C C G A A G T A A T A C A C T	a909_al2.seq
	C T T C A A T T T C T T T A T A G T C A A A T T C T T G C A T C A C T A A A T C T T C A C G G T T C	Majority
	13160 13170 13180 13190 13200	
13139	C T T C A A T T T C T T T A T A G T C A A A T T C T T G C A T A A C T A A A T C T T C A C G G T T C	cohl_al2.seq
13092	C T T C A A T T T C T T T A T A G T C A A A T T C T T G C A T C A C T A A A T C T T C A C G G T T C	a909_al2.seq
	A T A T C T T G A T A C A A A C A A G A T A A C A T A C C G A C C T T A G G T A A A T G A A G G T A	Majority
	13210 13220 13230 13240 13250	
13189	A T A T C T T G A T A C A A A C A A G A T A A C A T A C C G A C C T T A G G T A A A T G A A G G T A	cohl_al2.seq
13142	A T A T C T T G A T A C A A A C A A G A T A A C A T A C C G A C C T T A G G T A A A T G A A G G T A	a909_al2.seq
	A T T T T C A T A A T T A T C T A T C A A A T C A C C T A G C A C A C C G A A T C T T G A T C T A	Majority
	13260 13270 13280 13290 13300	
13239	A T T T T C A T A A T T A T C T A T C A A A T C A C C T A G C A C A C C G A A T C T T G A T C T A	cohl_al2.seq
13192	A T T T T C A T A A T T A T C T A T C A A A T C A C C T A G C A C A C C G A A T C T T G A T C T A	a909_al2.seq
	A A G T C A A G A A C C A A T C A A A A T T C T T G T G C T A C T G C A A A T T G A C C G A T A C A G	Majority
	13310 13320 13330 13340 13350	
13289	A A G T C A A G A A C C A A T C A A A A T T C T T G T G C T A C T G C A A A T T G A C C G A T A C A G	cohl_al2.seq
13242	A A G T C A A G A A C C A A T C A A A A T T C T T G T G C T A C T G C A A A T T G A C C A T A C A G	a909_al2.seq
	T T C A A A G C A T A T G C A A T T C C T T T A T T T T C T G T T A A A T A A T C A A C A G T T A G	Majority
	13360 13370 13380 13390 13400	
13339	T T C A A A G C A T A T G C A A T T C C T T T A T T T T C T G T T A A A T A A T C A A C A G T T A G	cohl_al2.seq
13292	T T C A A A G C A T A T G C A A T T C C T T T A T T T T C T G T T A A A T A A T C A A C A G T T A G	a909_al2.seq
	G T G C C C C T C T T C A T T A T A A T C G G C C A C T A A T T G A G A A A T T T C T T C C T T A T	Majority
	13410 13420 13430 13440 13450	
13389	G T G C C C C T C T T C A T T A T A A T C G G C C A C T A A T T G A G A A A T T T C T T C C T T A T	cohl_al2.seq
13342	G T G C C C C T C T T C A T T A T A A T C G G C C A C T A A T T G A G A A A T T T C T T C C T T A T	a909_al2.seq
	T T T T C G A G C C A T T A T C T A C C A T G T A G A T A T G G C T T A C T T G A G G A T A A A T T	Majority
	13460 13470 13480 13490 13500	
13439	T T T T C G A G C C A T T A T C T A C C A T G T A G A T A T G G C T T A C T T G A G G A T A A A T T	cohl_al2.seq
13392	T T T T C G A G C C A T T A T C T A C C A T G T A G A T A T G G C T T A C T T G A G G A T A A A T T	a909_al2.seq
	G C T C G A A T G T T C T G A T G T A A G C G T T C A A T A T T G C G G T T A A A G G T G A C A A T	Majority
	13510 13520 13530 13540 13550	
13489	G C T C G A A T G T T C T G A T G T A A G C G T T C A A T A T T G G G G T T A A A G G T G A C A A T	cohl_al2.seq
13442	G C T C G A A T G T T C T G A T G T A A G C G T T C A A T A T T G G G G T T A A A G G T G A C A A T	a909_al2.seq
	A C C C G C T A A A T A T T T C A T G T T C T A T G C T C T T T T C T A A A A T C T C T A A A T A A	Majority
	13560 13570 13580 13590 13600	
13539	A C C C G C T A A A T A T T T C A T G T T C T A T G C T C T T T T C T A A A A T C T C T A A A T A A	cohl_al2.seq
13492	A C C C G C T A A A T A T T T C A T G T T C T A T G C T C T T T T C T A A A A T C T C T A A A T A A	a909_al2.seq
	C T G A A T G A C T G G T G C T T T G G T T A T A A A A A C G A T A C C G A C A T A G A T A G T T A	Majority
	13610 13620 13630 13640 13650	
13589	C T G A A T G A C T G G T G C T T T G G T T A T A A A A A C G A T A C C G A C A T A G A T A G T T A	cohl_al2.seq
13542	C T G A A T G A C T G G T G C T T T G G T T A T A A A A A C G A T A C C G A C A T A G A T A G T T A	a909_al2.seq

FIGURE 21T

131/487

	CTGCTTACCTAAATCTTGAATGACATAATTTACCAATGATACTGACATTGTA	Majority
	13660 13670 13680 13690 13700	
13639	CTGCTACTAAACTTTTGAATGACATAATTTACCAATGATACTGACATTGTA	coh1_al2.seq
13592	CTGCTACTAAACTTTTGAATGACATAATTTACCAATGATACTGACATTGTA	a909_al2.seq
	GTATTGATATAATAGAGTACAGCTCCACTAAGAGTAGCAGCAATTAAATA	Majority
	13710 13720 13730 13740 13750	
13689	GTATTGATATAATAAGTACAGCTCCACTAAGAGTAGCAGCAATTAAATA	coh1_al2.seq
13642	GTATTGATATAATAGAGTACAGCTCCACTAAGAGTAGCAGCAATTAAATA	a909_al2.seq
	GCCGAGCATTTCCTCTTGTAAATTCTTTAAAGTAAATACATCTCTTAAAG	Majority
	13760 13770 13780 13790 13800	
13739	GCCGAGCATTTCCTCTTGTAAATTCTTTAAAGTAAATACATCTCTTAAAG	coh1_al2.seq
13692	GCCGAGCATTTCCTCTTGTAAATTCTTTAAAGTAAATACATCTCTTAAAG	a909_al2.seq
	AGATAGCTTTCATATAGGGAGACAATAAATTCAGTAATAACTGTAGAGATA	Majority
	13810 13820 13830 13840 13850	
13789	AGATAGCTTTCATATAGGGAGACAATAAATTCAGTAATAACTGTAGAGATA	coh1_al2.seq
13742	AGATAGCTTTCATATAGGGAGACAATAAATTCAGTAATAACTGTAGAGATA	a909_al2.seq
	ATAGCTCCCATAGCACCTAAATTTGGTATTAAAAAGTATATTAAGCACAAAC	Majority
	13860 13870 13880 13890 13900	
13839	ATAGCTCCCATAGCACCTAAATTTGGTATTAAAAAGTATATTAAGCACAAAC	coh1_al2.seq
13792	ATAGCTCCCATAGCACCTAAATTTGGTATTAAAAAGTATATTAAGCACAAAC	a909_al2.seq
	ATTTGCCACAAGTCCAATAACTGCAGACATTGTGTAAGCTTTTGTACGTC	Majority
	13910 13920 13930 13940 13950	
13889	ATTTGCCACAAGTCCAATAACTGCAGACATTGTGTAAGCTTTTGTACGTC	coh1_al2.seq
13842	ATTTGCCACAAGTCCAATAACTGCAGACATTGTGTAAGCTTTTGTACGTC	a909_al2.seq
	TTGAAGCCAGTAGATACTGTGTCCCTAAAGCGTTACCATAAGAAATGCAA	Majority
	13960 13970 13980 13990 14000	
13939	TTGAAGCCAGTAGATACTGTGTCCCTAAAGCGTTACCATAAGAAATGCAA	coh1_al2.seq
13892	TTGAAGCCAGTAGATACTGTGTCCCTAAAGCGTTACCATAAGAAATGCAA	a909_al2.seq
	ATGATCATCAAA	Majority
	14010	
13989	ATGATCATCAAA	coh1_al2.seq
13942	ATGATCATCAAA	a909_al2.seq

Decoration 'Decoration #1': Shade (with solid black) residues that differ from the Consensus.

FIGURE 21U

132/487

M K L S K K L L F S A A V L T M V A G S T V E P V A Q F A T G M S I V R A A E V S Q E R P A K T T V Majority									
10 20 30 40 50									
1	M	K	L	S	K	K	L	L	F
1	M	K	L	S	K	K	L	L	F
1	M	K	L	S	K	K	L	L	F
1	M	K	L	S	K	K	L	L	F
1	M	K	L	S	K	K	L	L	F
N I Y K L Q A D S Y K S E I T S N G G I E N K D G E V I S N Y A K L G D N V K G L Q C V Q F K R Y K Majority									
60 70 80 90 100									
51	N	I	Y	K	L	Q	A	D	S
51	N	I	Y	K	L	Q	A	D	S
51	N	I	Y	K	L	Q	A	D	S
51	N	I	Y	K	L	Q	A	D	S
51	N	I	Y	K	L	Q	A	D	S
V K T D I S V D E L K K L T T V E A A D A K V G T I L E E G V S L P Q K T N A Q C L V D A L D S K Majority									
110 120 130 140 150									
01	V	K	T	D	I	S	V	D	E
01	V	K	T	D	I	S	V	D	E
01	V	K	T	D	I	S	V	D	E
01	V	K	T	D	I	S	V	D	E
01	V	K	T	D	I	S	V	D	E
S N V R Y L Y V E D L K N S P S N I T K A Y A V P F V L E L P V A N S T G T G F L S E I N I Y P K N Majority									
160 170 180 190 200									
51	S	N	V	R	Y	L	Y	V	E
51	S	N	V	R	Y	L	Y	V	E
51	S	N	V	R	Y	L	Y	V	E
51	S	N	V	R	Y	L	Y	V	E
51	S	N	V	R	Y	L	Y	V	E
V V T D E P K T D K D V K K L G Q D D A G Y T I G E E F K W F L K S T I P A N L G D Y E K F E I T D Majority									
210 220 230 240 250									
01	V	V	T	D	E	P	K	T	D
01	V	V	T	D	E	P	K	T	D
01	V	V	T	D	E	P	K	T	D
01	V	V	T	D	E	P	K	T	D
01	V	V	T	D	E	P	K	T	D
K F A D G L T Y K S V G K I K I G S K T L N R D E H Y T I D E P T V D N Q N T L K I T F K P E K F K Majority									
260 270 280 290 300									
51	K	F	A	D	G	L	T	Y	K
51	K	F	A	D	G	L	T	Y	K
51	K	F	A	D	G	L	T	Y	K
51	K	F	A	D	G	L	T	Y	K
51	K	F	A	D	G	L	T	Y	K
E I A E L L K G M T L V K N Q D A L D K A T A N T D D A A F L E I P V A S T I N E K A V L G K A I E Majority									
310 320 330 340 350									
11	E	I	A	E	L	L	K	G	M
11	E	I	A	E	L	L	K	G	M
11	E	I	A	E	L	L	K	G	M
11	E	I	A	E	L	L	K	G	M
11	E	I	A	E	L	L	K	G	M
N T F E L Q Y D H T P D K A D N P K P S N P P R K P E V H T G G K R F V K K D S T E T Q T L G G A E Majority									
360 370 380 390 400									
11	N	T	F	E	L	Q	Y	D	H
11	N	T	F	E	L	Q	Y	D	H
11	N	T	F	E	L	Q	Y	D	H
11	N	T	F	E	L	Q	Y	D	H
11	N	T	F	E	L	Q	Y	D	H

Figure 22

133/487

FDLLASDGTAVKWTDALEIKANTNKNYIACEAVTCQPIKLSHTDGTFEIK Majority									
410		420		430		440		450	
01	F	D	L	L	A	S	D	G	T
01	F	D	L	L	A	S	D	G	T
01	F	D	L	L	A	S	D	G	T
01	F	D	L	L	A	S	D	G	T
01	F	D	L	L	A	S	D	G	T
GLAYAVDANAEGTAVTYKLKETKAPECYVIPDKEIEFTVSQTSYNTKPTD Majority									
460		470		480		490		500	
51	G	L	A	Y	A	V	D	A	N
51	G	L	A	Y	A	V	D	A	N
51	G	L	A	Y	A	V	D	A	N
51	G	L	A	Y	A	V	D	A	N
51	G	L	A	Y	A	V	D	A	N
ITVDSADATPDTIKNNKRPSIPNTGGIGTAIFVAIGA AVMAFAVKGMKRR Majority									
510		520		530		540		550	
01	I	T	V	D	S	A	D	A	T
01	I	T	V	D	S	A	D	A	T
01	I	T	V	D	S	A	D	A	T
01	I	T	V	D	S	A	D	A	T
01	I	T	V	D	S	A	D	A	T
TKDN Majority									
51	T	K	D	N					
51	T	K	D	N					
51	T	K	D	N					
51	T	K	D	N					
51	T	K	D	N					

scoration 'Decoration #1': Shade (with solid black) residues that differ from the Consensus.

FIGURE 22A

134/487

M K K R Q K I W R G L S V T L L I L S Q I P F G I L V Q G E T Q D T N Q A L															
---	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Figure 23

135/487

T S Y Q N Q F N S F L N K I P D R S G I L Q E D F I I N G D D Y Q I V K G D G E S F K L F S D R K V Majority															
410		420		430		440		450							
401	T S Y Q N Q F N S F L N K I P D R S G I L Q E D F I I N G D D Y Q I V K G D G E S F K L F S D R K V	sag649_2603.pep													
401	T S Y Q N Q F N S F L N K I P D R S G I L Q E D F I I N G D D Y Q I V K G D G E S F K L F S D R K V	sag649_coh1.pep													
401	T S Y Q N Q F N S F L N K I P D R S G I L Q E D F I I N G D D Y Q I V K G D G E S F K L F S D R K V	sag649_cjb111.pep													
401	T S Y Q N Q F N S F L N K I P D R S G I L Q E D F I I N G D D Y Q I V K G D G E S F K L F S D R K V	sag649_18rs21.pep													
401	T S Y Q N Q F N S F L N K I P D R S G I L Q E D F I I N G D D Y Q I V K G D G E S F K L F S D R K V	sag649_nem316.pep													
401	T S Y Q N Q F N S F L N K I P D R S G I L Q E D F I I N G D D Y Q I V K G D G E S F K L F S D R K V	sag649_cjb111.pep													
P V T G G T T Q A A Y R V P Q N Q L S V M S N E G Y A I N S G Y I Y L Y W R D Y N W V Y P F D P K T Majority															
460		470		480		490		500							
451	P V T G G T T Q A A Y R V P Q N Q L S V M S N E G Y A I N S G Y I Y L Y W R D Y N W V Y P F D P K T	sag649_2603.pep													
451	P V T G G T T Q A A Y R V P Q N Q L S V M S N E G Y A I N S G Y I Y L Y W R D Y N W V Y P F D P K T	sag649_coh1.pep													
451	P V T G G T T Q A A Y R V P Q N Q L S V M S N E G Y A I N S G Y I Y L Y W R D Y N W V Y P F D P K T	sag649_cjb111.pep													
451	P V T G G T T Q A A Y R V P Q N Q L S V M S N E G Y A I N S G Y I Y L Y W R D Y N W V Y P F D P K T	sag649_18rs21.pep													
451	P V T G G T T Q A A Y R V P Q N Q L S V M S N E G Y A I N S G Y I Y L Y W R D Y N W V Y P F D P K T	sag649_nem316.pep													
451	P V T G G T T Q A A Y R V P Q N Q L S V M S N E G Y A I N S G Y I Y L Y W R D Y N W V Y P F D P K T	sag649_cjb111.pep													
K K V S A T K Q I K T H G E P T T L Y F N G N I R P K G Y D I F T V G I G V N G D P G A T P L E A E Majority															
510		520		530		540		550							
501	K K V S A T K Q I K T H G E P T T L Y F N G N I R P K G Y D I F T V G I G V N G D P G A T P L E A E	sag649_2603.pep													
501	K K V S A T K Q I K T H G E P T T L Y F N G N I R P K G Y D I F T V G I G V N G D P G A T P L E A E	sag649_coh1.pep													
501	K K V S A T K Q I K T H G E P T T L Y F N G N I R P K G Y D I F T V G I G V N G D P G A T P L E A E	sag649_cjb111.pep													
501	K K V S A T K Q I K T H G E P T T L Y F N G N I R P K G Y D I F T V G I G V N G D P G A T P L E A E	sag649_18rs21.pep													
501	K K V S A T K Q I K T H G E P T T L Y F N G N I R P K G Y D I F T V G I G V N G D P G A T P L E A E	sag649_nem316.pep													
501	K K V S A T K Q I K T H G E P T T L Y F N G N I R P K G Y D I F T V G I G V N G D P G A T P L E A E	sag649_cjb111.pep													
K F M Q S I S S K T E N Y T N V D D T N K I Y D E L N K Y F K T I V E E K H S I V D G N V T D P M G Majority															
560		570		580		590		600							
551	K F M Q S I S S K T E N Y T N V D D T N K I Y D E L N K Y F K T I V E E K H S I V D G N V T D P M G	sag649_2603.pep													
551	K F M Q S I S S K T E N Y T N V D D T N K I Y D E L N K Y F K T I V E E K H S I V D G N V T D P M G	sag649_coh1.pep													
551	K F M Q S I S S K T E N Y T N V D D T N K I Y D E L N K Y F K T I V E E K H S I V D G N V T D P M G	sag649_cjb111.pep													
551	K F M Q S I S S K T E N Y T N V D D T N K I Y D E L N K Y F K T I V E E K H S I V D G N V T D P M G	sag649_18rs21.pep													
551	K F M Q S I S S K T E N Y T N V D D T N K I Y D E L N K Y F K T I V E E K H S I V D G N V T D P M G	sag649_nem316.pep													
551	K F M Q S I S S K T E N Y T N V D D T N K I Y D E L N K Y F K T I V E E K H S I V D G N V T D P M G	sag649_cjb111.pep													
E M I E F Q L K N G Q S F T H D D Y V L V G N D G S Q L K N G V A L G G P N S D G G I L K D V T V T Majority															
610		620		630		640		650							
601	E M I E F Q L K N G Q S F T H D D Y V L V G N D G S Q L K N G V A L G G P N S D G G I L K D V T V T	sag649_2603.pep													
601	E M I E F Q L K N G Q S F T H D D Y V L V G N D G S Q L K N G V A L G G P N S D G G I L K D V T V T	sag649_coh1.pep													
601	E M I E F Q L K N G Q S F T H D D Y V L V G N D G S Q L K N G V A L G G P N S D G G I L K D V T V T	sag649_cjb111.pep													
601	E M I E F Q L K N G Q S F T H D D Y V L V G N D G S Q L K N G V A L G G P N S D G G I L K D V T V T	sag649_18rs21.pep													
601	E M I E F Q L K N G Q S F T H D D Y V L V G N D G S Q L K N G V A L G G P N S D G G I L K D V T V T	sag649_nem316.pep													
601	E M I E F Q L K N G Q S F T H D D Y V L V G N D G S Q L K N G V A L G G P N S D G G I L K D V T V T	sag649_cjb111.pep													
Y D K T S Q T I K I N H L N L G S G Q K V V L T Y D V R L K D N Y I S N K F Y N T N N R T T L S P K Majority															
660		670		680		690		700							
651	Y D K T S Q T I K I N H L N L G S G Q K V V L T Y D V R L K D N Y I S N K F Y N T N N R T T L S P K	sag649_2603.pep													
651	Y D K T S Q T I K I N H L N L G S G Q K V V L T Y D V R L K D N Y I S N K F Y N T N N R T T L S P K	sag649_coh1.pep													
651	Y D K T S Q T I K I N H L N L G S G Q K V V L T Y D V R L K D N Y I S N K F Y N T N N R T T L S P K	sag649_cjb111.pep													
651	Y D K T S Q T I K I N H L N L G S G Q K V V L T Y D V R L K D N Y I S N K F Y N T N N R T T L S P K	sag649_18rs21.pep													
651	Y D K T S Q T I K I N H L N L G S G Q K V V L T Y D V R L K D N Y I S N K F Y N T N N R T T L S P K	sag649_nem316.pep													
651	Y D K T S Q T I K I N H L N L G S G Q K V V L T Y D V R L K D N Y I S N K F Y N T N N R T T L S P K	sag649_cjb111.pep													
S E K E P N T I R D F P I P K I R D V R E F P V L T I S N Q K K M G E V E F I K V N K D K H S E S L Majority															
710															

136/487

	810	820	830	840	850	Majority
	SPDGYIEVKTTPVVTFTIQNGEVTNLKADPNANKNQIGYLEGNGKHLITN					
801	SPDGYIEVKTTPVVTFTIQNGEVTNLKADPNANKNQIGYLEGNGKHLITN					sag649_2603.pep
801	SPDGYIEVKTTPVVTFTIQNGEVTNLKADPNANKNQIGYLEGNGKHLITN					sag649_coh1.pep
801	SPDGYIEVKTTPVVTFTIQNGEVTNLKADPNANKNQIGYLEGNGKHLITN					sag649_cjb111.pep
801	SPDGYIEVKTTPVVTFTIQNGEVTNLKADPNANKNQIGYLEGNGKHLITN					sag649_18rs21.pep
801	SPDGYIEVKTTPVVTFTIQNGEVTNLKADPNANKNQIGYLEGNGKHLITN					sag649_new316.pep
801	SPDGYIEVKTTPVVTFTIQNGEVTNLKADPNANKNQIGYLEGNGKHLITN					sag649_cjb111.pep
	TPKRPPGCVFPKTCGGIGTIVYILVGSTFMILTICSFRRKQL					Majority
	860	870	880	890		
851	TPKRPPGCVFPKTCGGIGTIVYILVGSTFMILTICSFRRKQL					sag649_2603.pep
851	TPKRPPGCVFPKTCGGIGTIVYILVGSTFMILTICSFRRKQL					sag649_coh1.pep
851	TPKRPPGCVFPKTCGGIGTIVYILVGSTFMILTICSFRRKQL					sag649_cjb111.pep
851	TPKRPPGCVFPKTCGGIGTIVYILVGSTFMILTICSFRRKQL					sag649_18rs21.pep
851	TPKRPPGCVFPKTCGGIGTIVYILVGSTFMILTICSFRRKQL					sag649_new316.pep
851	TPKRPPGCVFPKTCGGIGTIVYILVGSTFMILTICSFRRKQL					sag649_cjb111.pep

Decoration 'Decoration #1': Shade (with solid black) residues that differ from the Consensus.

FIGURE 23B

137/487

MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGA K G K L V V K K T D D Q Majority				
10	20	30	40	50
1	MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGA K G K L V V K K T D D Q	sag1408_2603.pep		
1	MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGA K G K L V V K K T D D Q	sag1408_515.pep		
1	MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGA K G K L V V K K T D D Q	sag1408_18rs21.seq		
1	MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGA K G K L V V K K T D D Q	sag1408_cjb111.pep		
1	MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGA K G K L V V K K T D D Q	sag1408_h36b.pep		
1	MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGA K G K L V V K K T D D Q	sag1408_nem316.pep		
NKPLSKATFVLKTTAHPE SK IEKVTAELTGEATFDNLIPG D Y T L S E E T A P Majority				
60	70	80	90	100
51	NKPLSKATFVLKTTAHPE SK IEKVTAELTGEATFDNLIPG D Y T L S E E T A P	sag1408_2603.pep		
51	NKPLSKATFVLKTTAHPE SK IEKVTAELTGEATFDNLIPG D Y T L S E E T A P	sag1408_515.pep		
51	NKPLSKATFVLKTTAHPE SK IEKVTAELTGEATFDNLIPG D Y T L S E E T A P	sag1408_18rs21.seq		
51	NKPLSKATFVLKTTAHPE SK IEKVTAELTGEATFDNLIPG D Y T L S E E T A P	sag1408_cjb111.pep		
51	NKPLSKATFVLKTTAHPE SK IEKVTAELTGEATFDNLIPG D Y T L S E E T A P	sag1408_h36b.pep		
51	NKPLSKATFVLKTTAHPE SK IEKVTAELTGEATFDNLIPG D Y T L S E E T A P	sag1408_nem316.pep		
EGYKKTNTQTWQVKVESNGKTTIQNSGDKNSTIGQNQEELDKQY P P T G I Y E Majority				
110	120	130	140	150
101	EGYKKTNTQTWQVKVESNGKTTIQNSGDKNSTIGQNQEELDKQY P P T G I Y E	sag1408_2603.pep		
101	EGYKKTNTQTWQVKVESNGKTTIQNSGDKNSTIGQNQEELDKQY P P T G I Y E	sag1408_515.pep		
101	EGYKKTNTQTWQVKVESNGKTTIQNSGDKNSTIGQNQEELDKQY P P T G I Y E	sag1408_18rs21.seq		
101	EGYKKTNTQTWQVKVESNGKTTIQNSGDKNSTIGQNQEELDKQY P P T G I Y E	sag1408_cjb111.pep		
101	EGYKKTNTQTWQVKVESNGKTTIQNSGDKNSTIGQNQEELDKQY P P T G I Y E	sag1408_h36b.pep		
101	EGYKKTNTQTWQVKVESNGKTTIQNSGDKNSTIGQNQEELDKQY P P T G I Y E	sag1408_nem316.pep		
DTKESYKLEHVKG S V P N G K S E A K A V N P Y S S E G E H I R E I P E G T L S K R I S E V Majority				
160	170	180	190	200
151	DTKESYKLEHVKG S V P N G K S E A K A V N P Y S S E G E H I R E I P E G T L S K R I S E V	sag1408_2603.pep		
151	DTKESYKLEHVKG S V P N G K S E A K A V N P Y S S E G E H I R E I P E G T L S K R I S E V	sag1408_515.pep		
151	DTKESYKLEHVKG S V P N G K S E A K A V N P Y S S E G E H I R E I P E G T L S K R I S E V	sag1408_18rs21.seq		
151	DTKESYKLEHVKG S V P N G K S E A K A V N P Y S S E G E H I R E I P E G T L S K R I S E V	sag1408_cjb111.pep		
151	DTKESYKLEHVKG S V P N G K S E A K A V N P Y S S E G E H I R E I P E G T L S K R I S E V	sag1408_h36b.pep		
151	DTKESYKLEHVKG S V P N G K S E A K A V N P Y S S E G E H I R E I P E G T L S K R I S E V	sag1408_nem316.pep		
GDLA H N K Y K I E L T V S G K T I V K P V D K Q K P L D V V F V L D N S N S M N N D G P N F Q R Majority				
210	220	230	240	250
201	GDLA H N K Y K I E L T V S G K T I V K P V D K Q K P L D V V F V L D N S N S M N N D G P N F Q R	sag1408_2603.pep		
201	GDLA H N K Y K I E L T V S G K T I V K P V D K Q K P L D V V F V L D N S N S M N N D G P N F Q R	sag1408_515.pep		
201	GDLA H N K Y K I E L T V S G K T I V K P V D K Q K P L D V V F V L D N S N S M N N D G P N F Q R	sag1408_18rs21.seq		
201	GDLA H N K Y K I E L T V S G K T I V K P V D K Q K P L D V V F V L D N S N S M N N D G P N F Q R	sag1408_cjb111.pep		
201	GDLA H N K Y K I E L T V S G K T I V K P V D K Q K P L D V V F V L D N S N S M N N D G P N F Q R	sag1408_h36b.pep		
201	GDLA H N K Y K I E L T V S G K T I V K P V D K Q K P L D V V F V L D N S N S M N N D G P N F Q R	sag1408_nem316.pep		
H N K A K K A A E A L G T A V K D I L G A N S D N R V A L V T Y G S D I F D G R S V D V V K G F K E Majority				
260	270	280	290	300
251	H N K A K K A A E A L G T A V K D I L G A N S D N R V A L V T Y G S D I F D G R S V D V V K G F K E	sag1408_2603.pep		
251	H N K A K K A A E A L G T A V K D I L G A N S D N R V A L V T Y G S D I F D G R S V D V V K G F K E	sag1408_515.pep		
251	H N K A K K A A E A L G T A V K D I L G A N S D N R V A L V T Y G S D I F D G R S V D V V K G F K E	sag1408_18rs21.seq		
251	H N K A K K A A E A L G T A V K D I L G A N S D N R V A L V T Y G S D I F D G R S V D V V K G F K E	sag1408_cjb111.pep		
247	H N K A K K A A E A L G T A V K D I L G A N S D N R V A L V T Y G S D I F D G R S V D V V K G F K E	sag1408_h36b.pep		
251	H N K A K K A A E A L G T A V K D I L G A N S D N R V A L V T Y G S D I F D G R S V D V V K G F K E	sag1408_nem316.pep		
D D K Y Y G L Q T K F T I Q T E N Y S H K Q L T N N A E E I I K R I P T E A P K A K W G S T T N G L Majority				
310	320	330	340	350
301	D D K Y Y G L Q T K F T I Q T E N Y S H K Q L T N N A E E I I K R I P T E A P K A K W G S T T N G L	sag1408_2603.pep		
301	D D K Y Y G L Q T K F T I Q T E N Y S H K Q L T N N A E E I I K R I P T E A P K A K W G S T T N G L	sag1408_515.pep		
301	D D K Y Y G L Q T K F T I Q T E N Y S H K Q L T N N A E E I I K R I P T E A P K A K W G S T T N G L	sag1408_18rs21.seq		
301	D D K Y Y G L Q T K F T I Q T E N Y S H K Q L T N N A E E I I K R I P T E A P K A K W G S T T N G L	sag1408_cjb111.pep		
297	D D K Y Y G L Q T K F T I Q T E N Y S H K Q L T N N A E E I I K R I P T E A P K A K W G S T T N G L	sag1408_h36b.pep		
301	D D K Y Y G L Q T K F T I Q T E N Y S H K Q L T N N A E E I I K R I P T E A P K A K W G S T T N G L	sag1408_nem316.pep		
T P E Q Q K E Y Y L S K V G E T F T M K A F M E A D D I L S Q V N R N S Q K I I V H V T D G V P T R Majority				
360	370	380	390	400
351	T P E Q Q K E Y Y L S K V G E T F T M K A F M E A D D I L S Q V N R N S Q K I I V H V T D G V P T R	sag1408_2603.pep		
351	T P E Q Q K E Y Y L S K V G E T F T M K A F M E A D D I L S Q V N R N S Q K I I V H V T D G V P T R	sag1408_515.pep		
351	T P E Q Q K E Y Y L S K V G E T F T M K A F M E A D D I L S Q V N R N S Q K I I V H V T D G V P T R	sag1408_18rs21.seq		
351	T P E Q Q K E Y Y L S K V G E T F T M K A F M E A D D I L S Q V N R N S Q K I I V H V T D G V P T R	sag1408_cjb111.pep		
346	T P E Q Q K E Y Y L S K V G E T F T M K A F M E A D D I L S Q V N R N S Q K I I V H V T D G V P T R	sag1408_h36b.pep		
351	T P E Q Q K E Y Y L S K V G E T F T M K A F M E A D D I L S Q V N R N S Q K I I V H V T D G V P T R	sag1408_nem316.pep		

Figure 24

138/487

SYAINNFKLGCASYESQFEQMKKNGYLNKSNFLLTDKPEDIKNGGESYFLF Majority									
401	SYA	I	N	N	F	K	L	G	C
401	SYA	I	N	N	F	K	L	G	C
401	SYA	I	N	N	F	K	L	G	C
401	SYA	I	N	N	F	K	L	G	C
396	SYA	I	N	N	F	K	L	G	C
401	SYA	I	N	N	F	K	L	G	C
PLDSYQTQIIISGNLQKLHYLDLNLNYPKGTIYRNGPVEHGTPTKLYINS Majority									
451	PLD	S	Y	Q	T	Q	I	I	S
451	PLD	S	Y	Q	T	Q	I	I	S
451	PLD	S	Y	Q	T	Q	I	I	S
451	PLD	S	Y	Q	T	Q	I	I	S
446	PLD	S	Y	Q	T	Q	I	I	S
451	PLD	S	Y	Q	T	Q	I	I	S
LKQKNYDIFNFGIDISGFRQVYNEFYKKNQDGTFFQKLKEEAFELSDGEIT Majority									
501	LKQ	K	N	Y	D	I	F	N	F
501	LKQ	K	N	Y	D	I	F	N	F
501	LKQ	K	N	Y	D	I	F	N	F
501	LKQ	K	N	Y	D	I	F	N	F
496	LKQ	K	N	Y	D	I	F	N	F
501	LKQ	K	N	Y	D	I	F	N	F
ELMRSFSSSKPEYYTPIVTSADTSNNEILSKIQQQFETILTKEINSIVNGTI Majority									
551	ELM	R	S	F	S	S	K	P	E
551	ELM	R	S	F	S	S	K	P	E
551	ELM	R	S	F	S	S	K	P	E
551	ELM	R	S	F	S	S	K	P	E
546	ELM	R	S	F	S	S	K	P	E
551	ELM	R	S	F	S	S	K	P	E
EDPMGDKINLQLGNGQTLQPSDYTLQGNDSVMKDGCIATGCPNNDGGILK Majority									
601	EDP	M	G	D	K	I	N	L	Q
601	EDP	M	G	D	K	I	N	L	Q
601	EDP	M	G	D	K	I	N	L	Q
601	EDP	M	G	D	K	I	N	L	Q
596	EDP	M	G	D	K	I	N	L	Q
601	EDP	M	G	D	K	I	N	L	Q
GVKLEYIGNKLYVRGLNLGEGQKVTLTYDVKLDDSFISNKFYDTNGRTTL Majority									
651	G	V	K	L	E	Y	I	G	N
651	G	V	K	L	E	Y	I	G	N
651	G	V	K	L	E	Y	I	G	N
651	G	V	K	L	E	Y	I	G	N
646	G	V	K	L	E	Y	I	G	N
651	G	V	K	L	E	Y	I	G	N
NPKSEDPNTLRDFFIPKIRDVREYPTITIKNEKKLGEIEFIKVDKDNKKL Majority									
701	N	P	K	S	E	D	P	N	T
701	N	P	K	S	E	D	P	N	T
701	N	P	K	S	E	D	P	N	T
701	N	P	K	S	E	D	P	N	T
696	N	P	K	S	E	D	P	N	T
701	N	P	K	S	E	D	P	N	T
LLKGATFELQEFNEDYKLYLPIKNNNSKVVTGENGKISYKDLKDGKYQLI Majority									
751	L	L	K	G	A	T	F	E	L
751	L	L	K	G	A	T	F	E	L
751	L	L	K	G	A	T	F	E	L
751	L	L	K	G	A	T	F	E	L
746	L	L	K	G	A	T	F	E	L
751	L	L	K	G	A	T	F	E	L

FIGURE 24A

139/487

E A V S P E D Y Q K I T N K P I L T F E V V K G S I Q N I I A V N K Q I S E Y H E E G D K H L I T N Majority										
	810		820		830		840		850	
801	E A V S P E D Y Q K I T N K P I L T F E V V K G S I Q N I I A V N K Q I S E Y H E E G D K H L I T N	sag1408_2603.pep								
801	E A V S P E D Y Q K I T N K P I L T F E V V K G S I Q N I I A V N K Q I S E Y H E E G D K H L I T N	sag1408_515.pep								
801	E A V S P E D Y Q K I T N K P I L T F E V V K G S I Q N I I A V N K Q I S E Y H E E G D K H L I T N	sag1408_18rs21.seq								
796	E A V S P E D Y Q K I T N K P I L T F E V V K G S I Q N I I A V N K Q I S E Y H E E G D K H L I T N	sag1408_cjb111.pep								
801	E A V S P E D Y Q K I T N K P I L T F E V V K G S I Q N I I A V N K Q I S E Y H E E G D K H L I T N	sag1408_h36b.pep								
801	E A V S P E D Y Q K I T N K P I L T F E V V K G S I Q N I I A V N K Q I S E Y H E E G D K H L I T N	sag1408_nen316.pep								

T H I P P K G I I P M T G G K G I L S F I L I G G A M M S I A G G I Y I W K R Y K K S S D M S I E K Majority										
	860		870		880		890		900	
851	T H I P P K G I I P M T G G K G I L S F I L I G G A M M S I A G G I Y I W K R Y K K S S D M S I E K	sag1408_2603.pep								
851	T H I P P K G I I P M T G G K G I L S F I L I G G A M M S I A G G I Y I W K R Y K K S S D M S I E K	sag1408_515.pep								
851	T H I P P K G I I P M T G G K G I L S F I L I G G A M M S I A G G I Y I W K R Y K K S S D M S I E K	sag1408_18rs21.seq								
851	T H I P P K G I I P M T G G K G I L S F I L I G G A M M S I A G G I Y I W K R Y K K S S D M S I E K	sag1408_cjb111.pep								
846	T H I P P K G I I P M T G G K G I L S F I L I G G A M M S I A G G I Y I W K R Y K K S S D M S I E K	sag1408_h36b.pep								
851	T H I P P K G I I P M T G G K G I L S F I L I G G A M M S I A G G I Y I W K R Y K K S S D M S I E K	sag1408_nen316.pep								

D

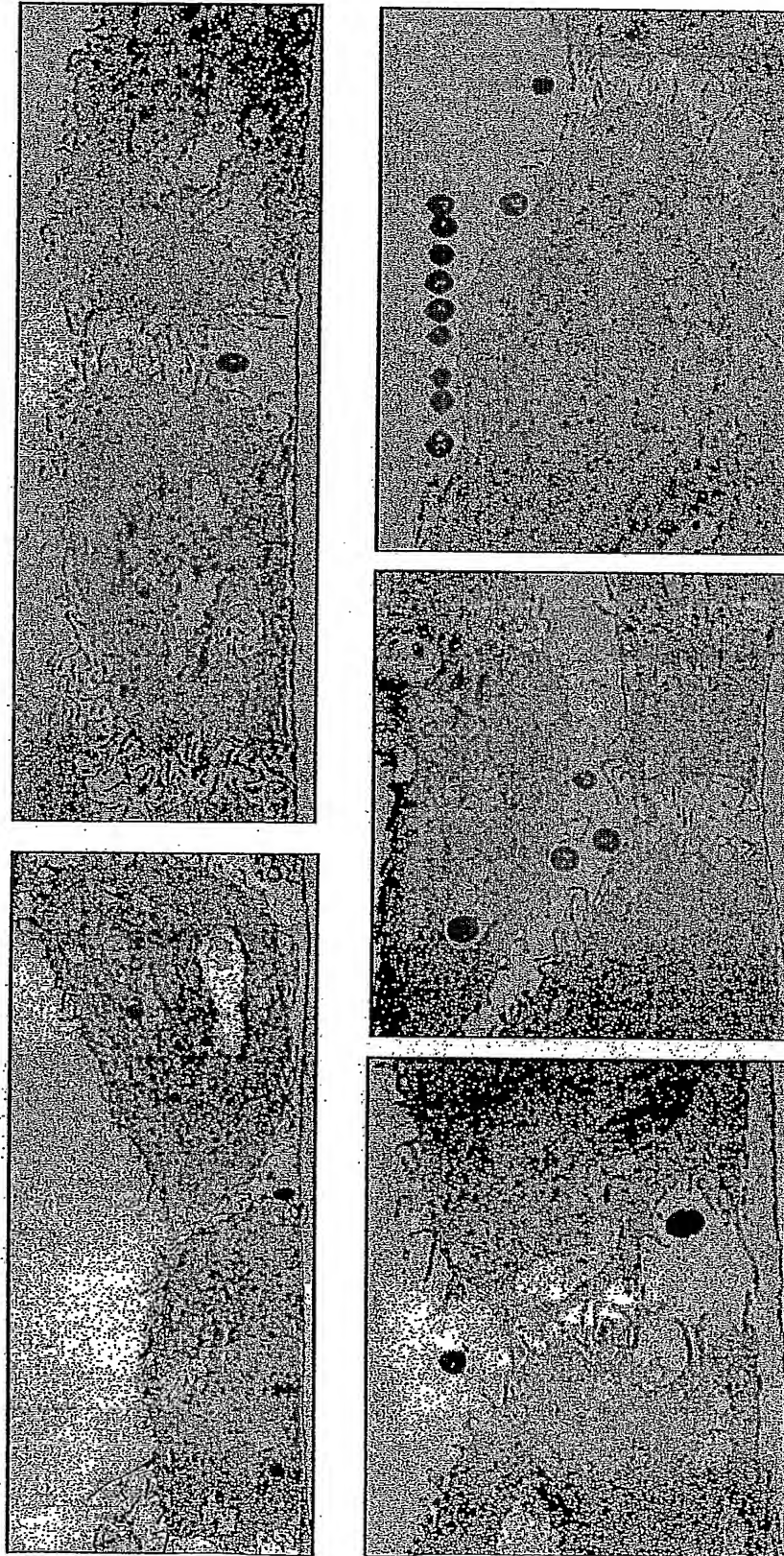
Majority

901 D	sag1408_2603.pep
901 D	sag1408_515.pep
901 D	sag1408_18rs21.seq
901 D	sag1408_cjb111.pep
896 D	sag1408_h36b.pep
901 D	sag1408_nen316.pep

Decoration 'Decoration #1': Shade (with solid black) residues that differ from the Consensus.

FIGURE 24B

Figure 25: GBS closely associate with tight junctions and cross the monolayer by a paracellular route



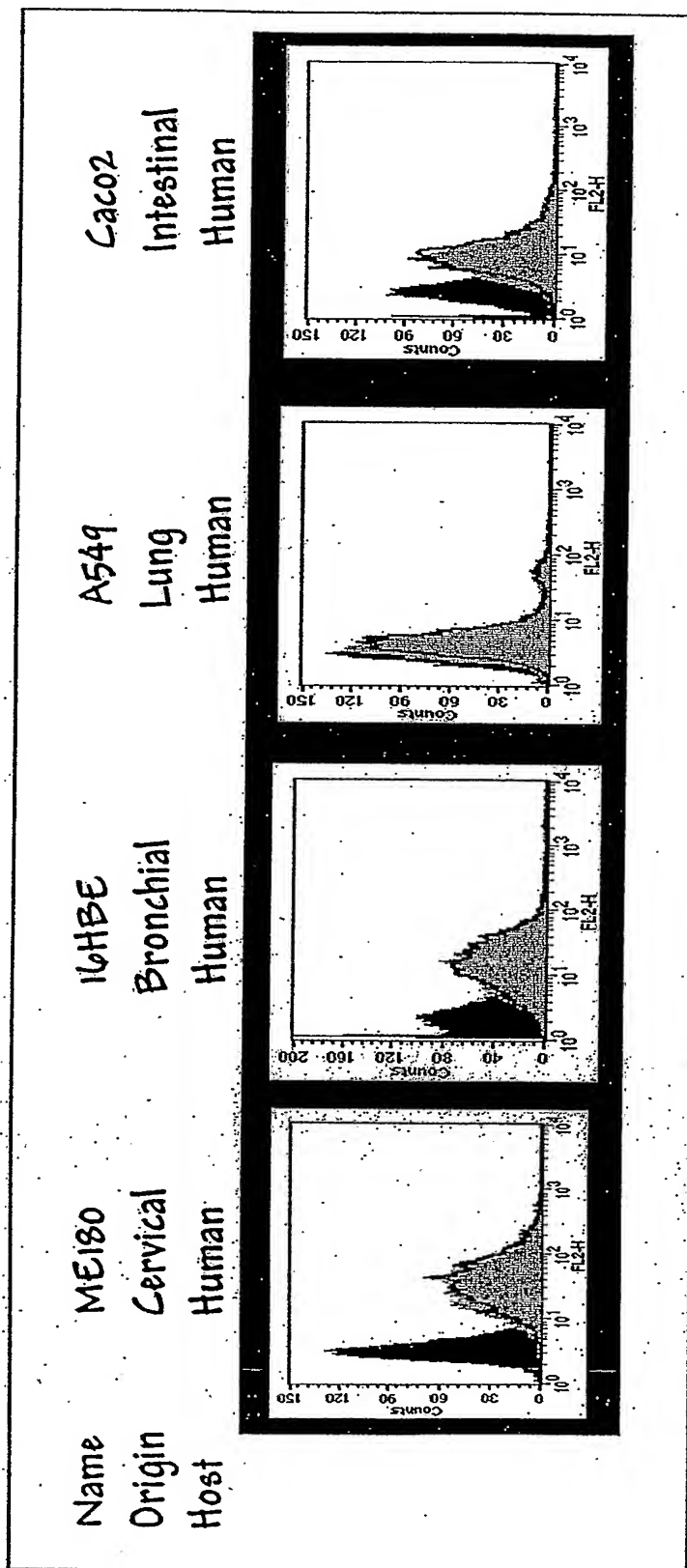
Transmission Electron Microscopy images of GBS infection of ME180 cervical epithelial cells.

Figure 26: GBS infection of ME180 cells
GBS infection of ME180 cells



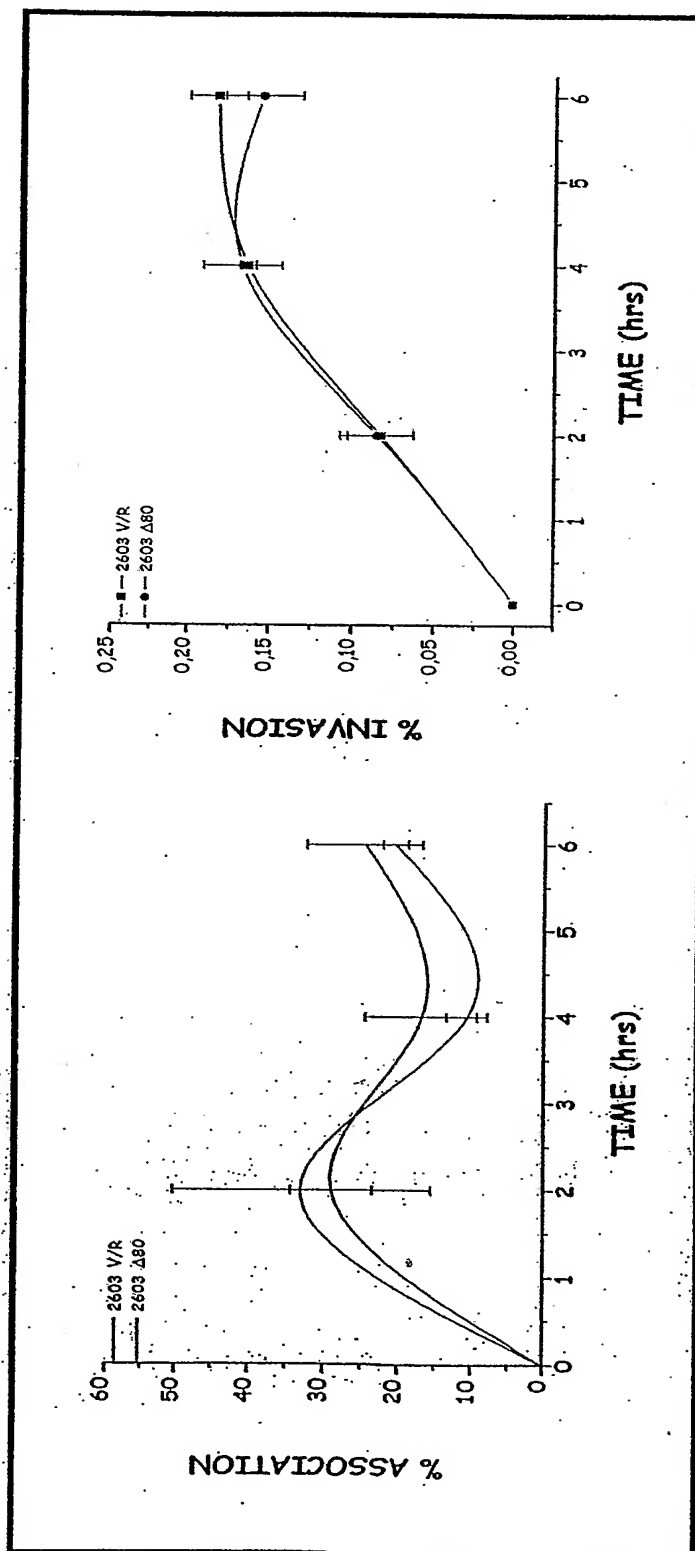
PCT/US05/27239/42/487

Figure 27



PCT/US05/27239 143/487

Figure 28



144/487

Figure 29

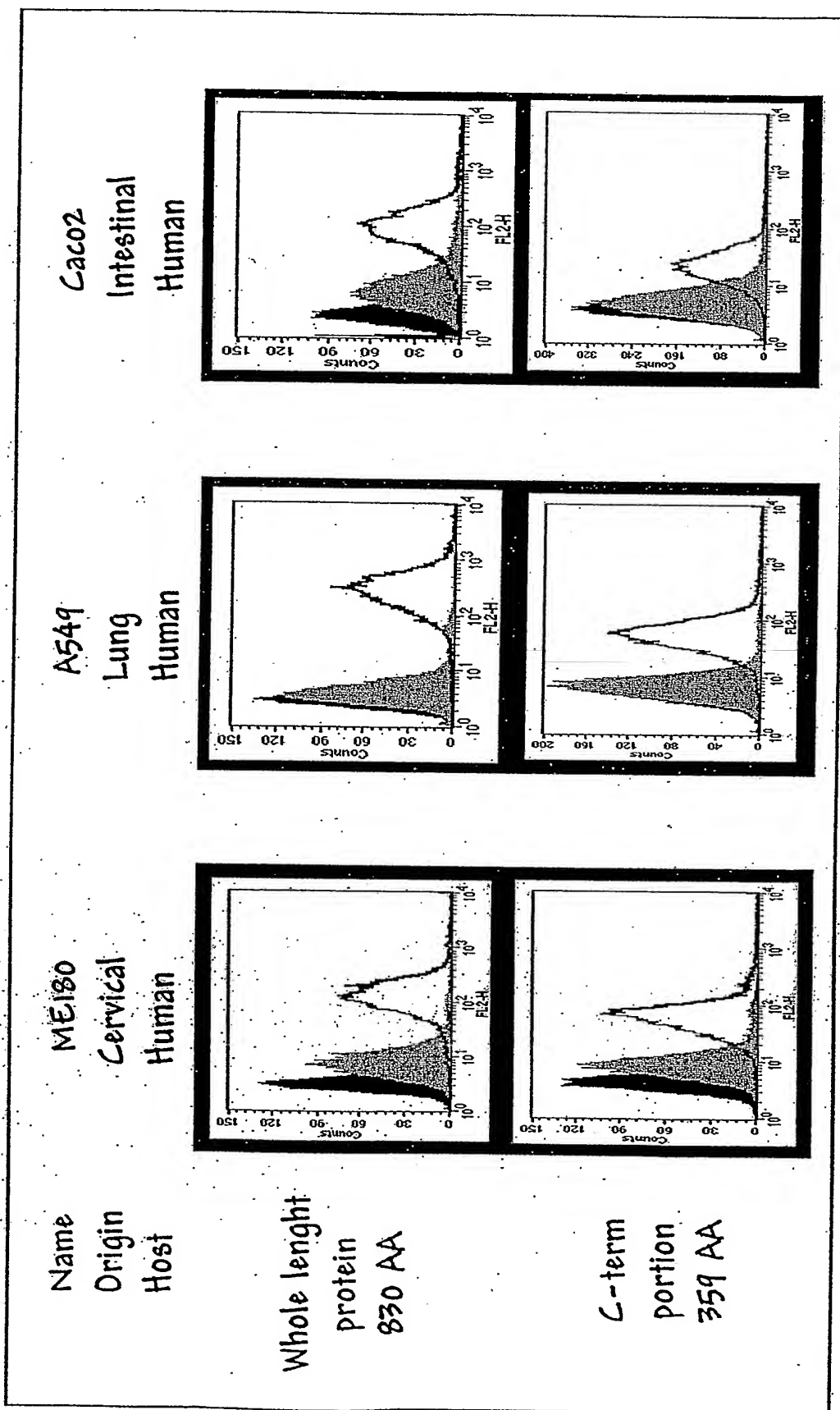
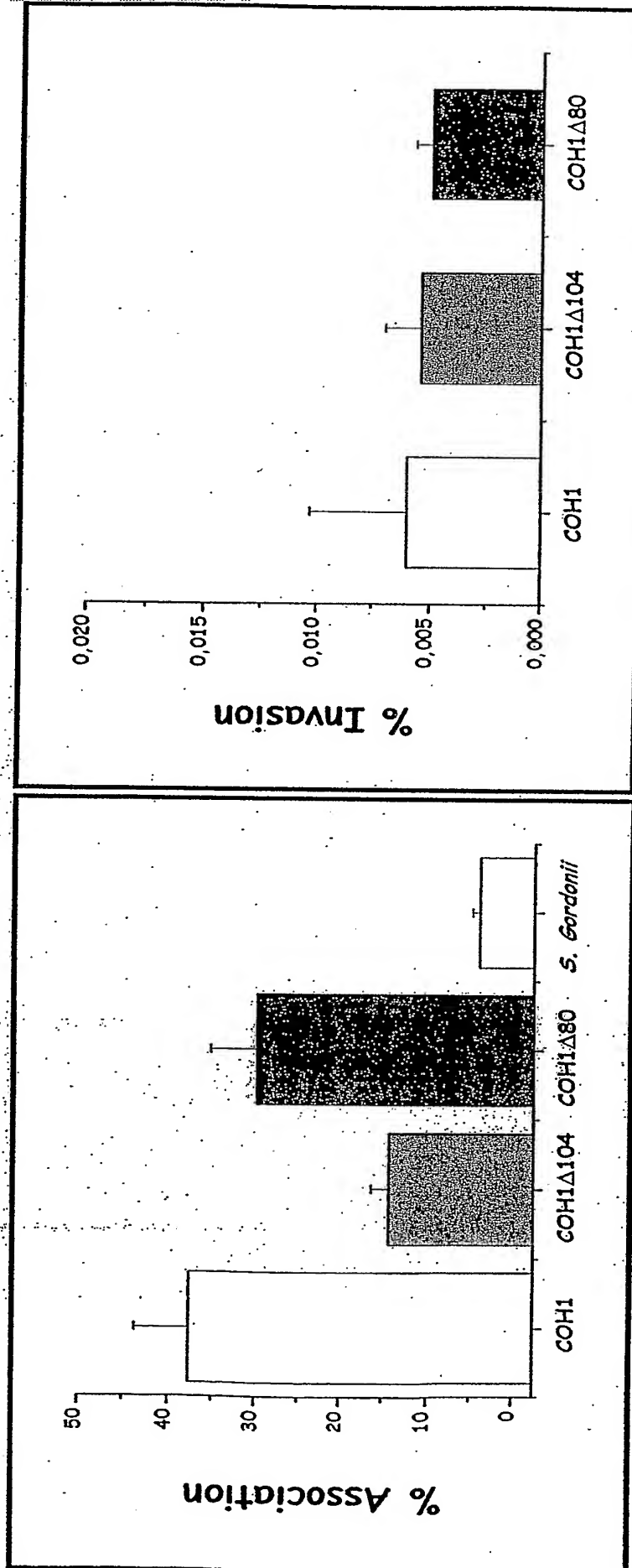
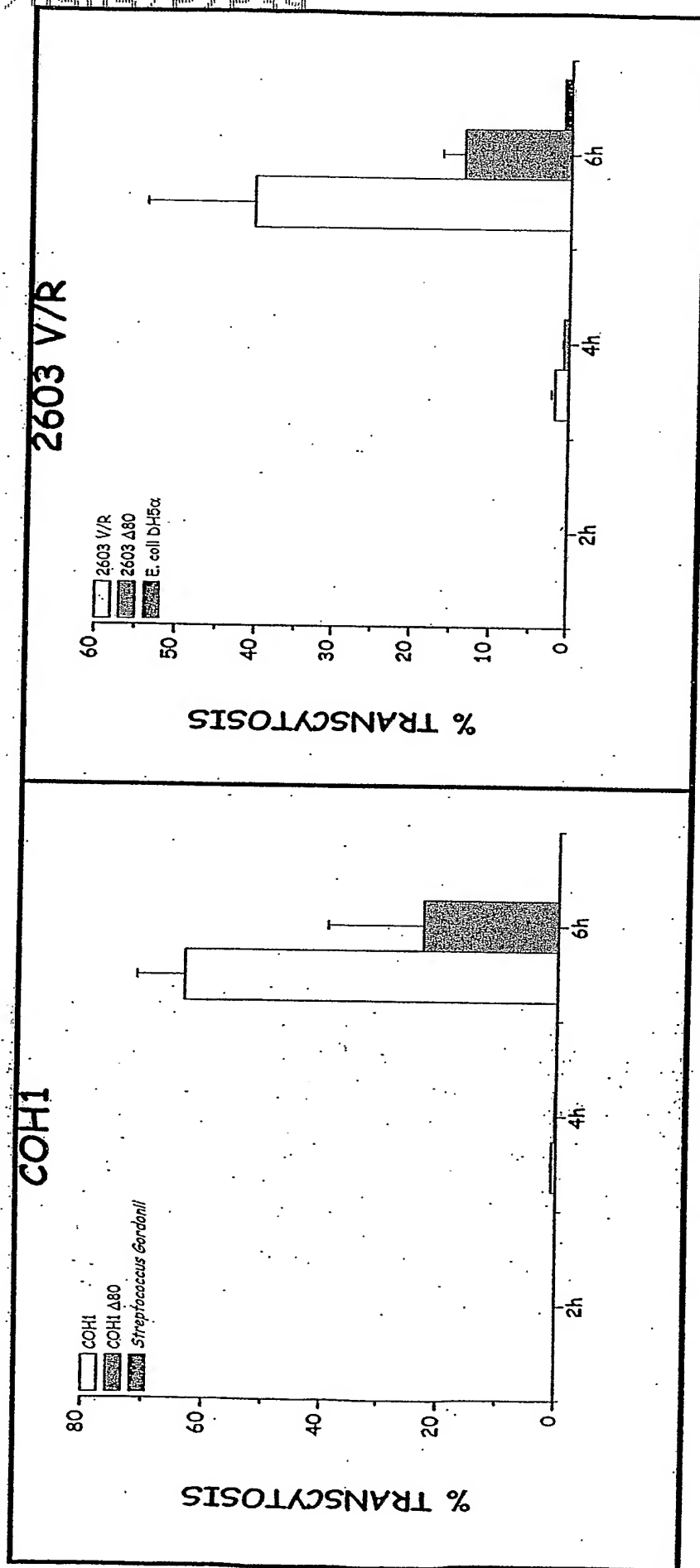


Figure 30



PCT/US2005/027239

Figure 31



PCT/US2005/027239/47/487

Figure 32

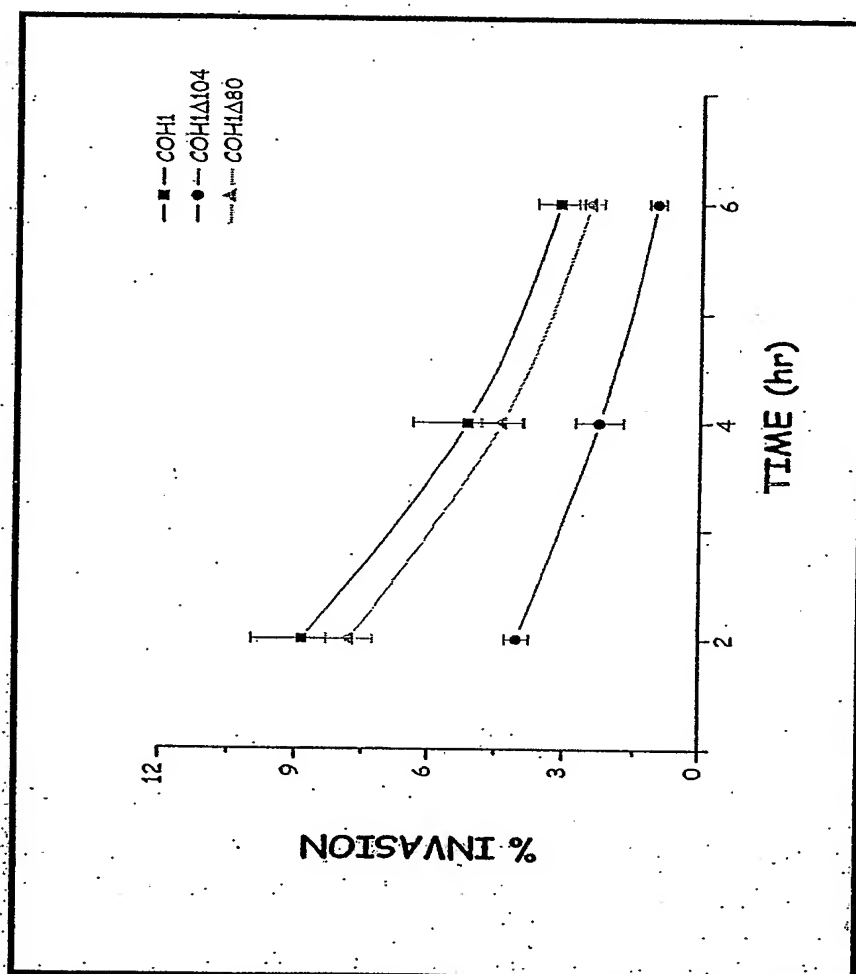
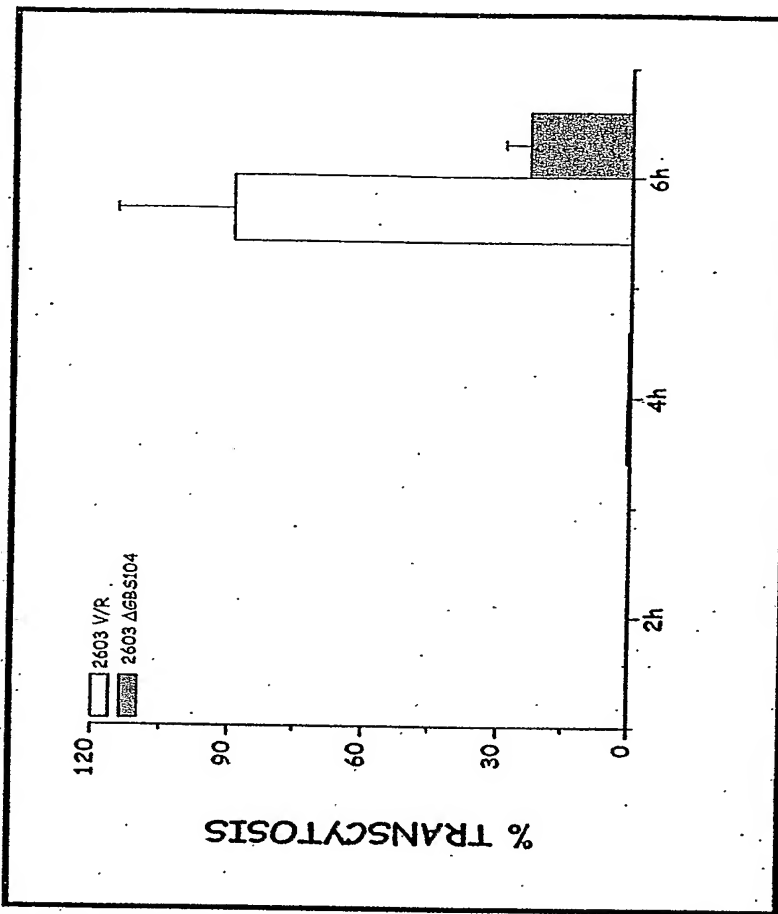
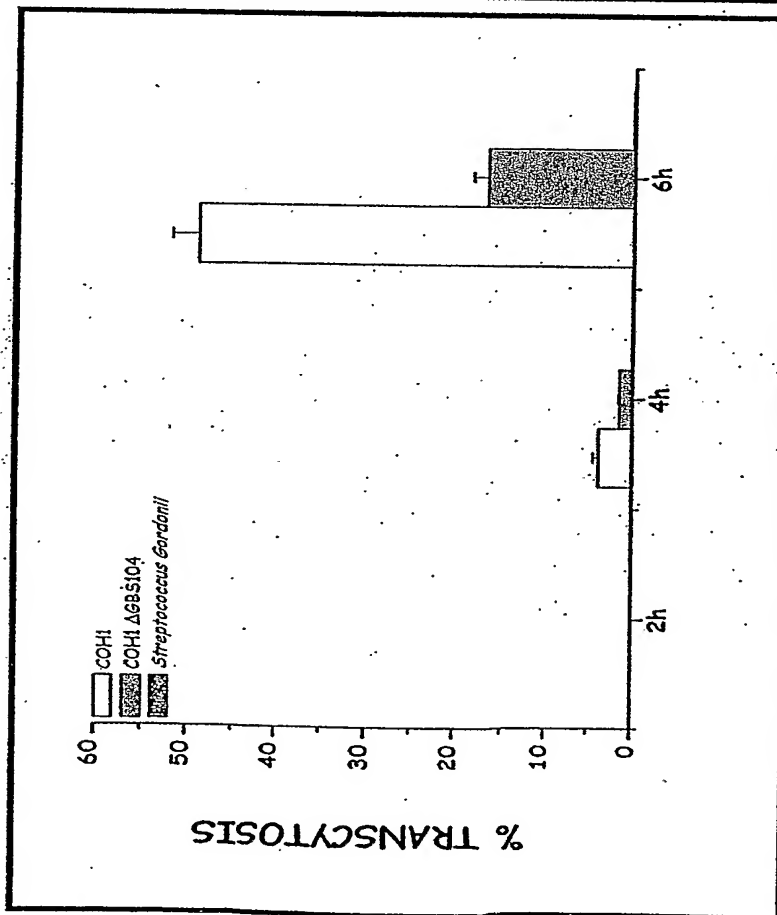


Figure 33

COH1

2603 V/R

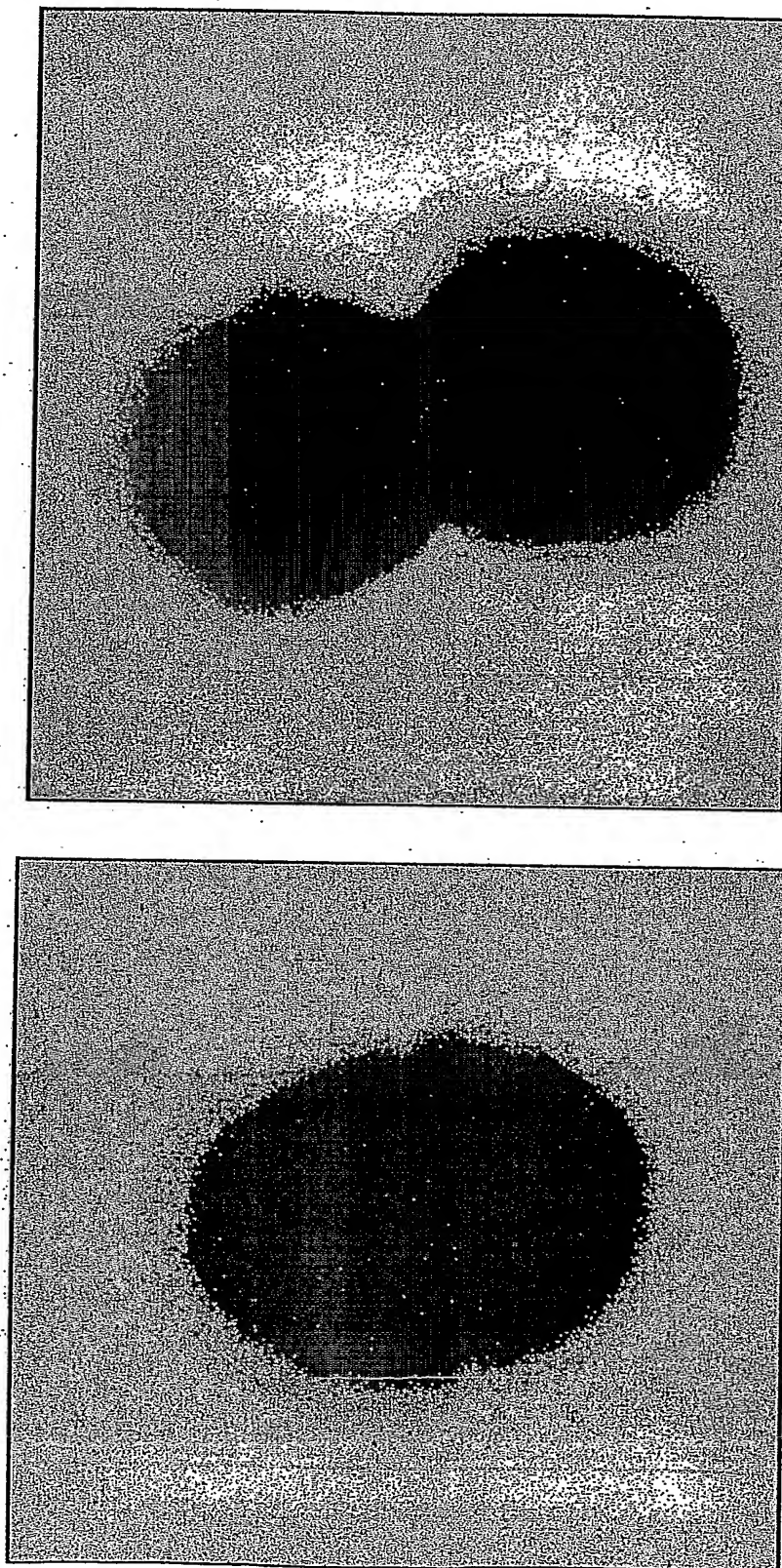


PCT/US05/27239 149/487

GBS STRAIN COH1 over GBS80

Figure 34

Negative staining EM



GBS STRAIN COH1 over GBS80

IEM anti-GBS80 (gold particles 10nm)

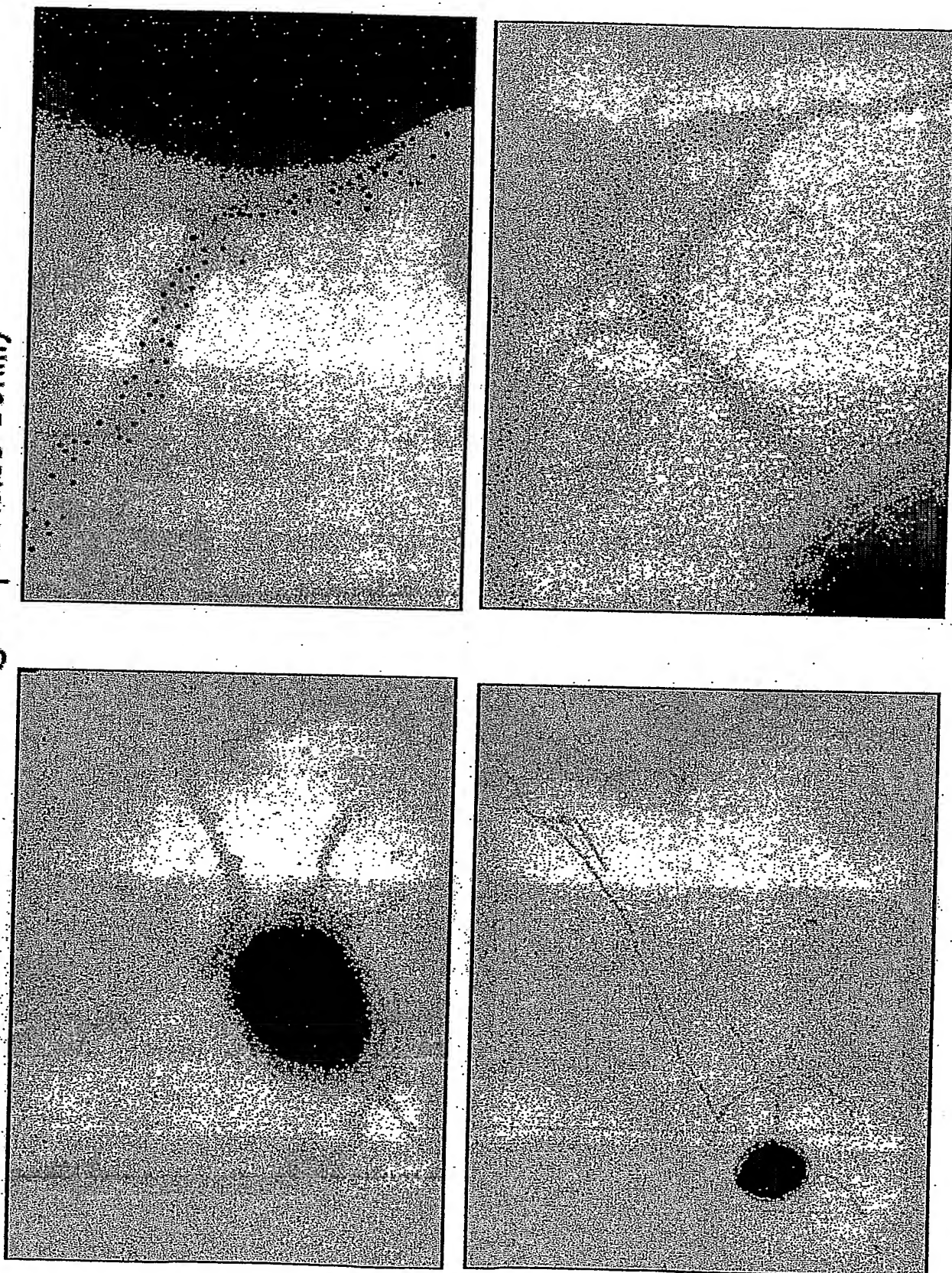
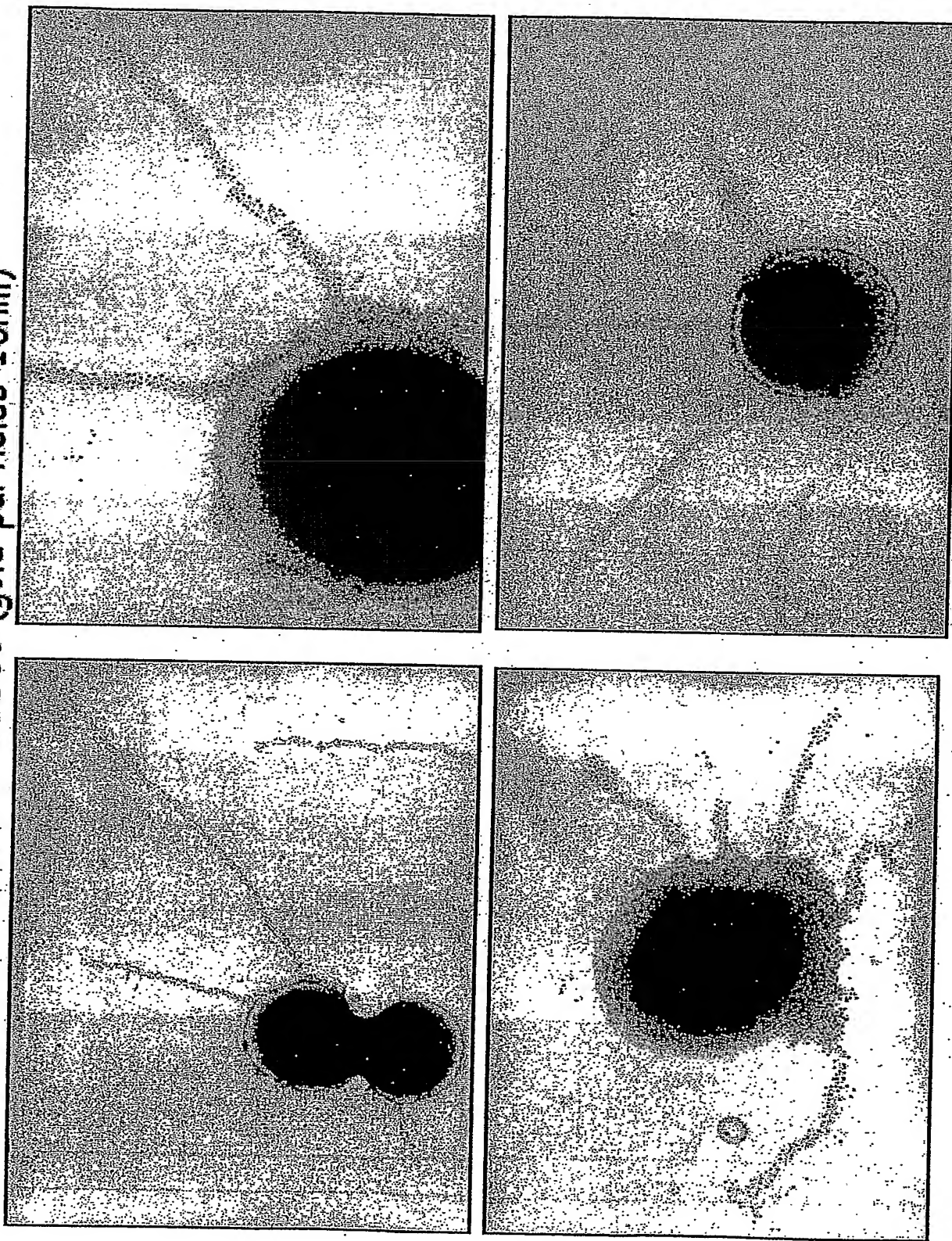


Figure 35

GBS STRAIN COH1 over GBS80

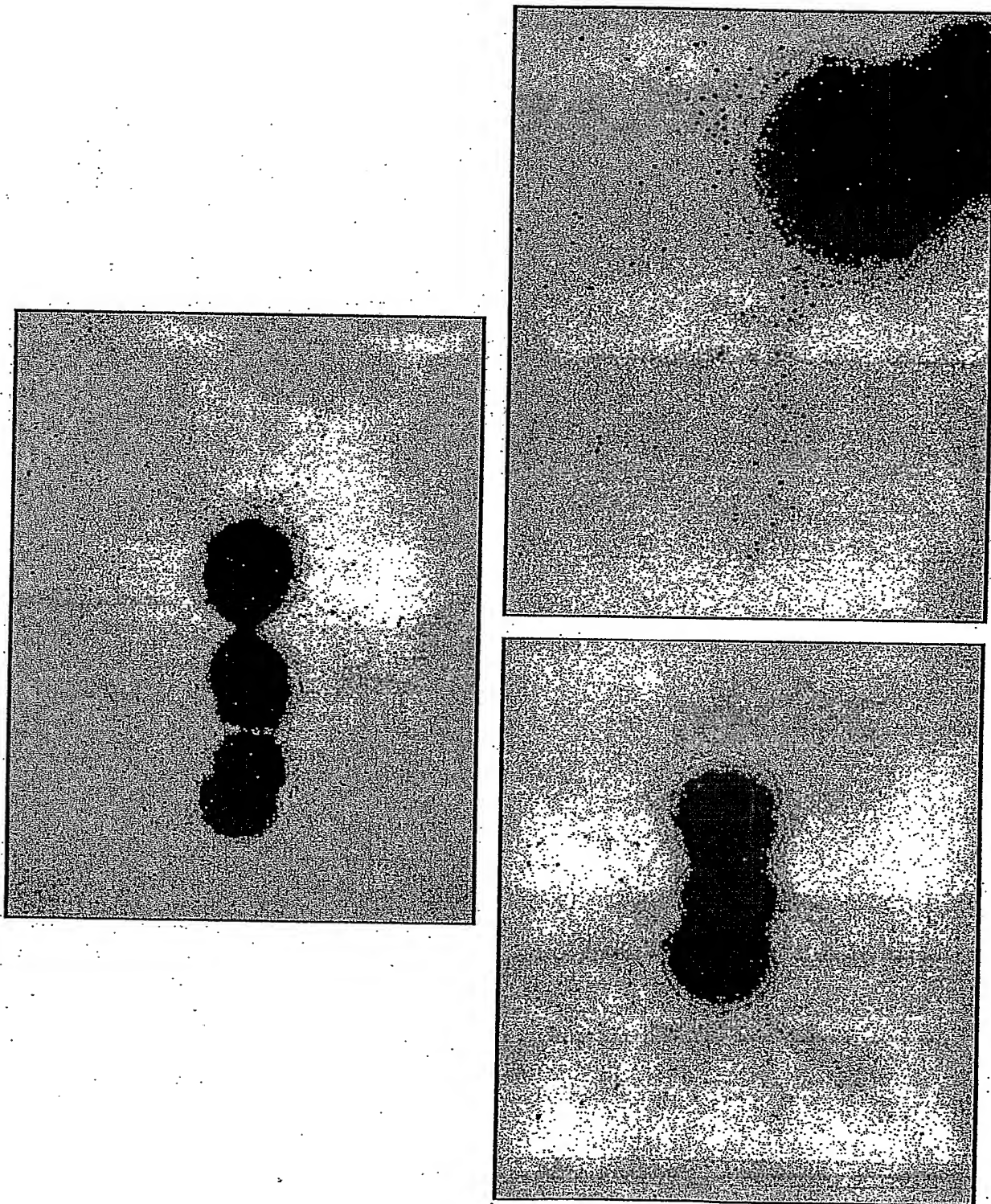
Figure 36

IEM anti-GBS80 (gold particles 10nm)



GBS STRAIN COH1 over GBS80
IEM anti-GBS80 (gold particles 20nm)

Figure 37



GBS STRAIN COH1 over GBS80

IEM anti-GBS104 (gold particles 10nm)

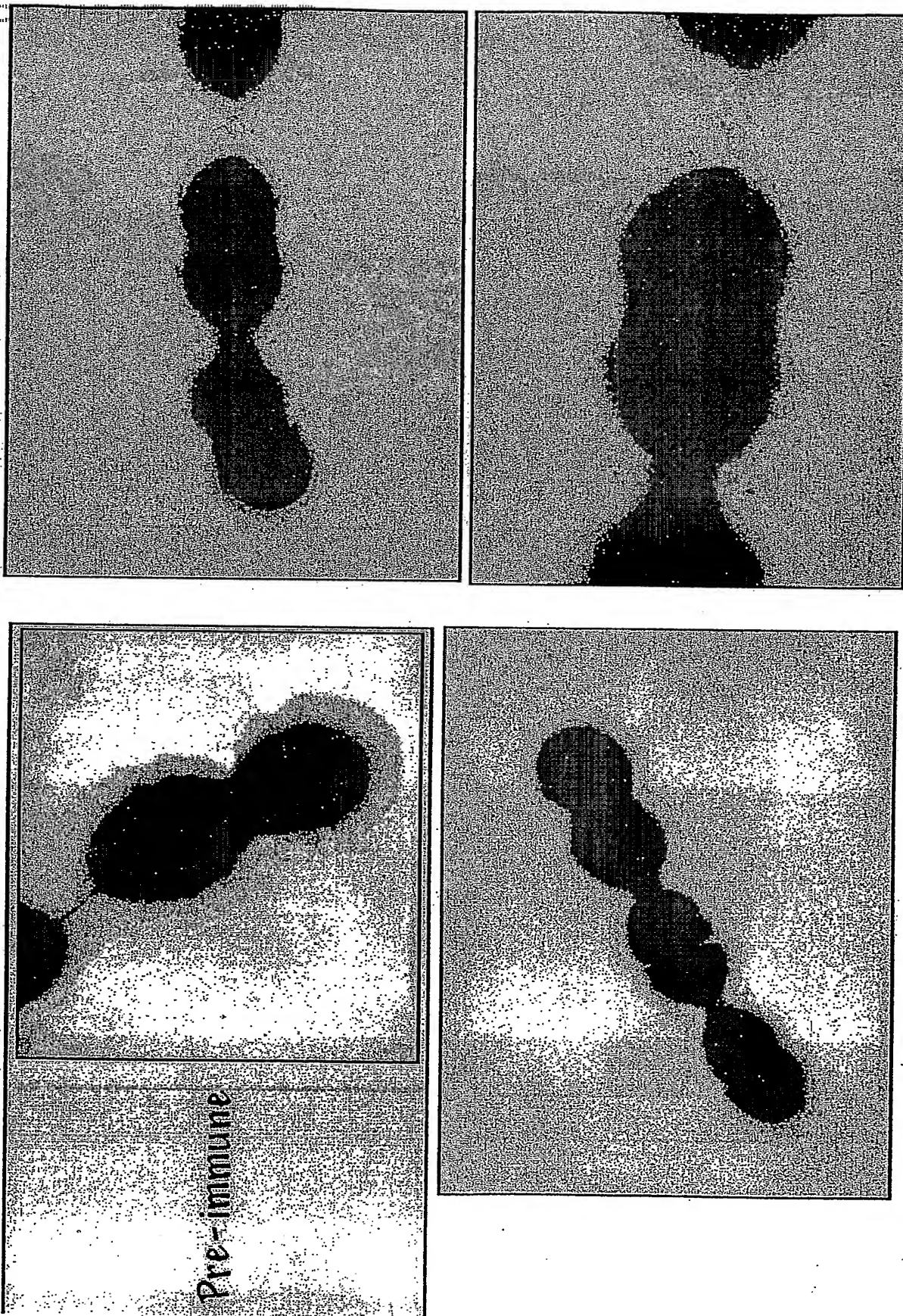


Figure 38

GBS STRAIN COH1 over GBS80

Figure 39

IEM anti-GBS80 (gold particles 20nm) anti-GBS104 (gold particles 10nm)

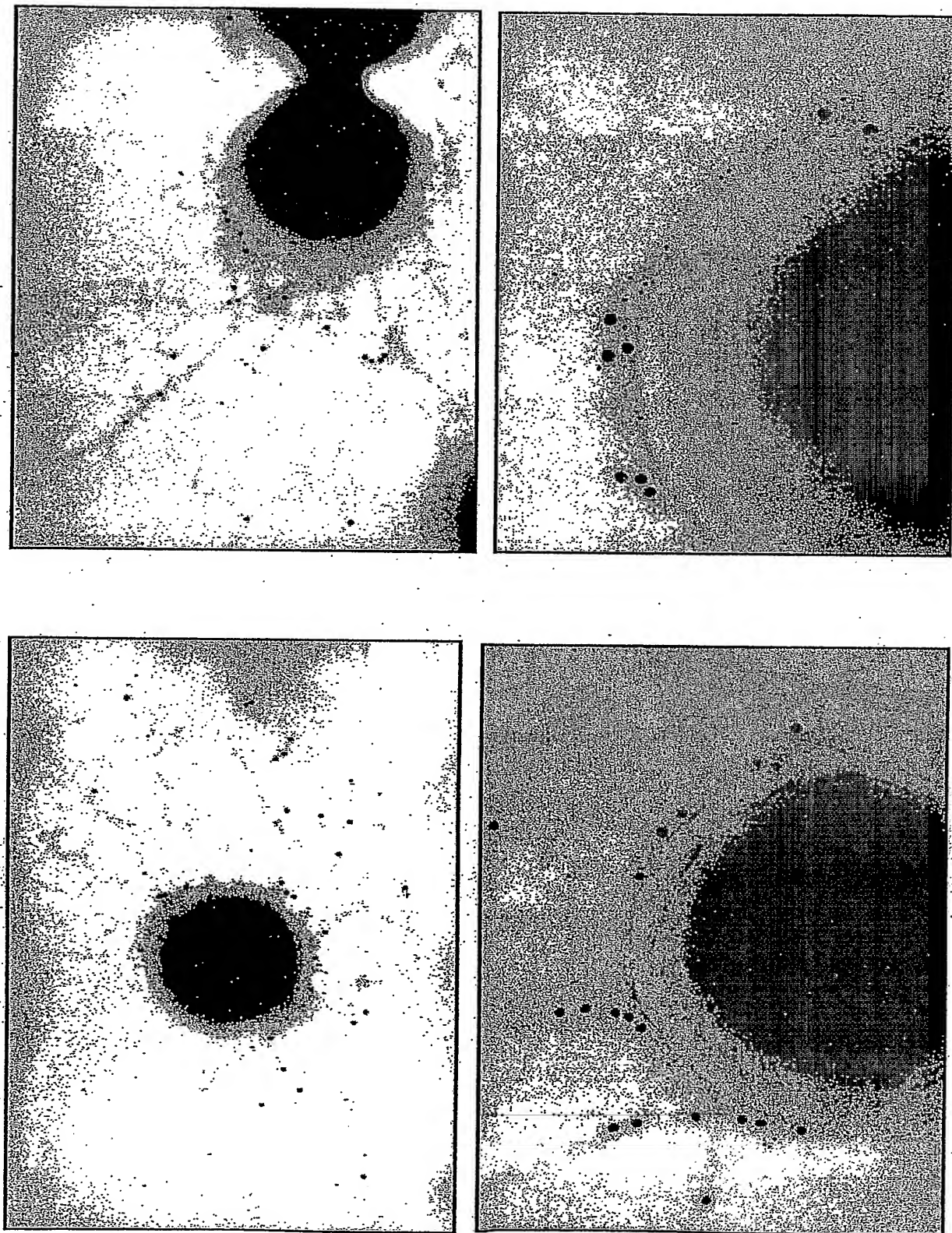
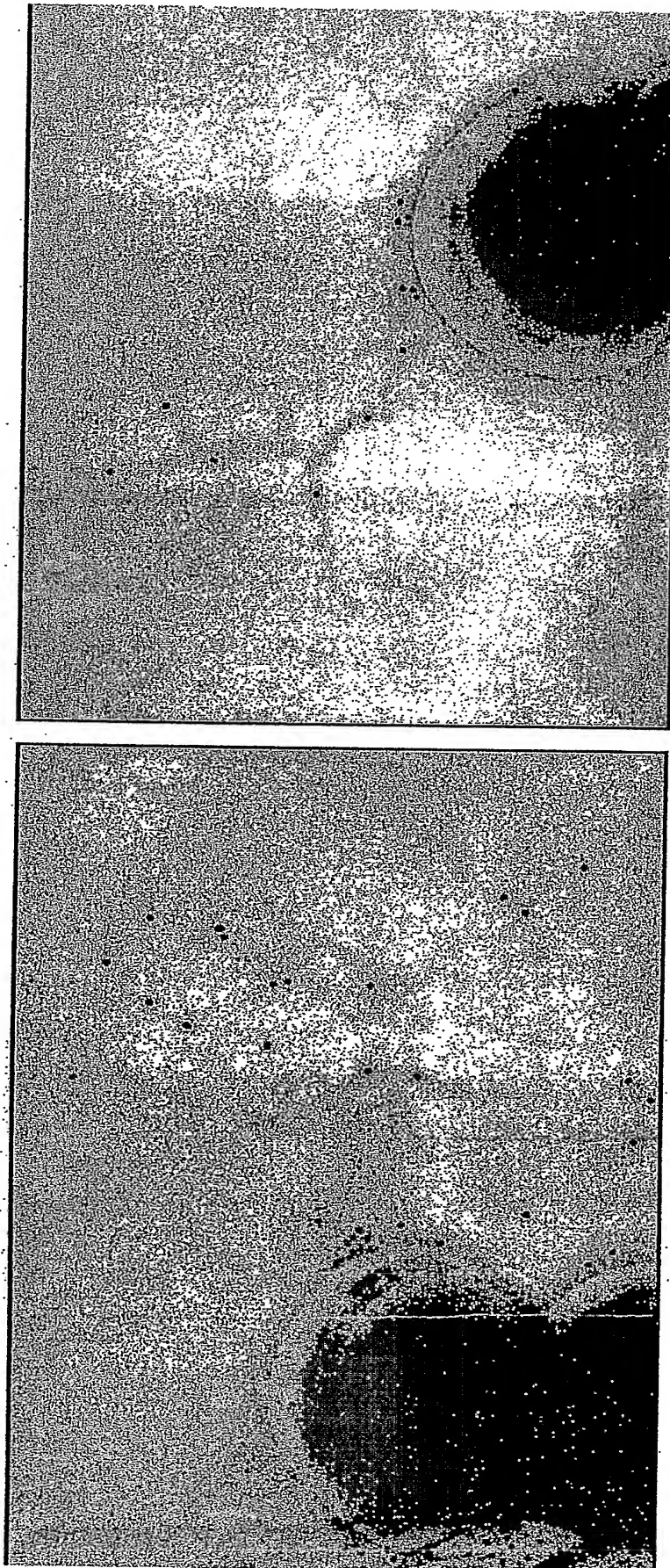
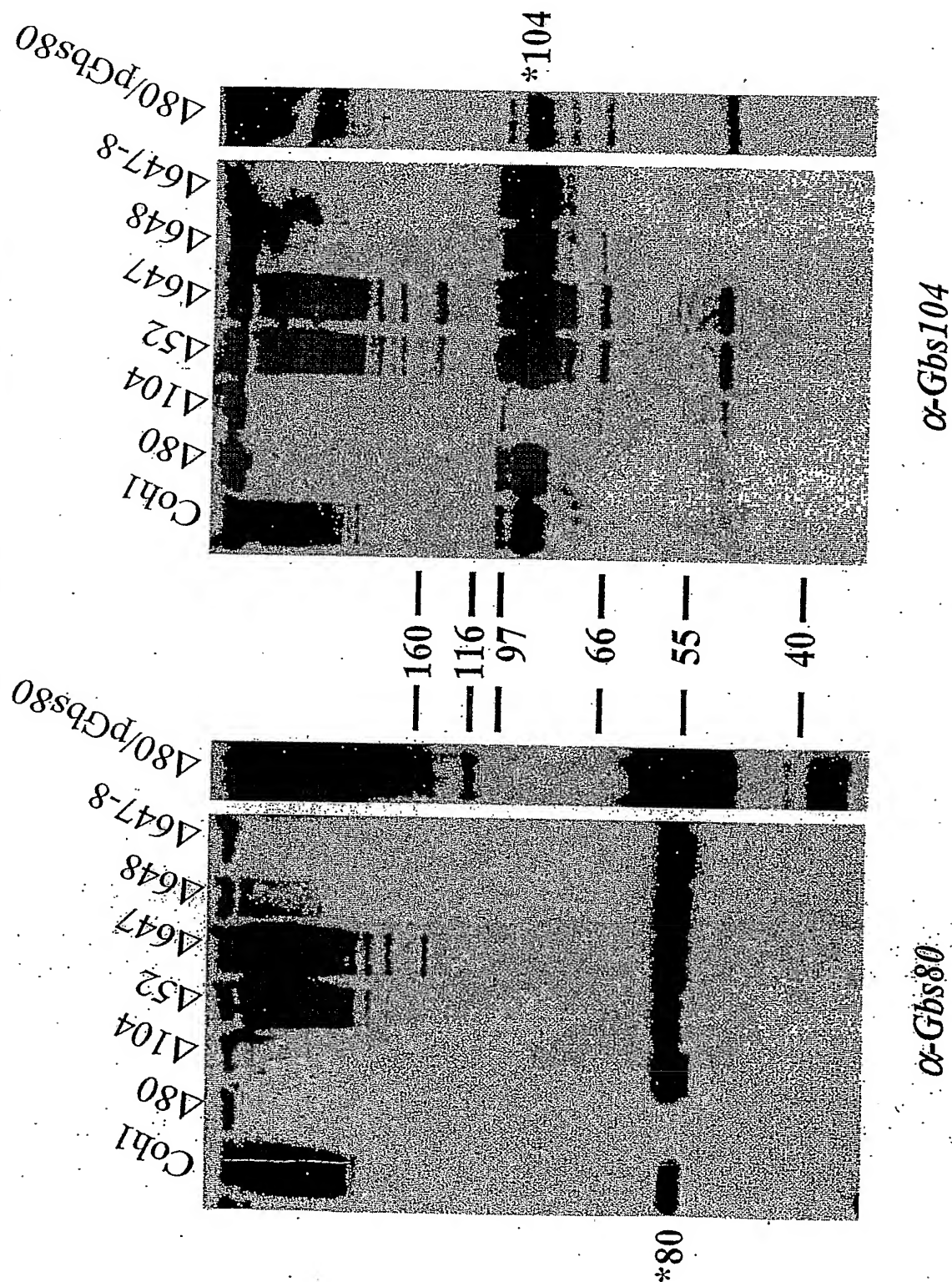


Figure 40

GBS STRAIN COH1 over GBS80**IEM anti-GBS80 (gold particles 20nm) anti-GBS104 (gold particles 10nm)**

156/487

Figure 41: GBS 80 is necessary for polymer formation, GBS104 and sortase SAG0648 are necessary for efficient assembly



PCT/US2005/027239 157/487

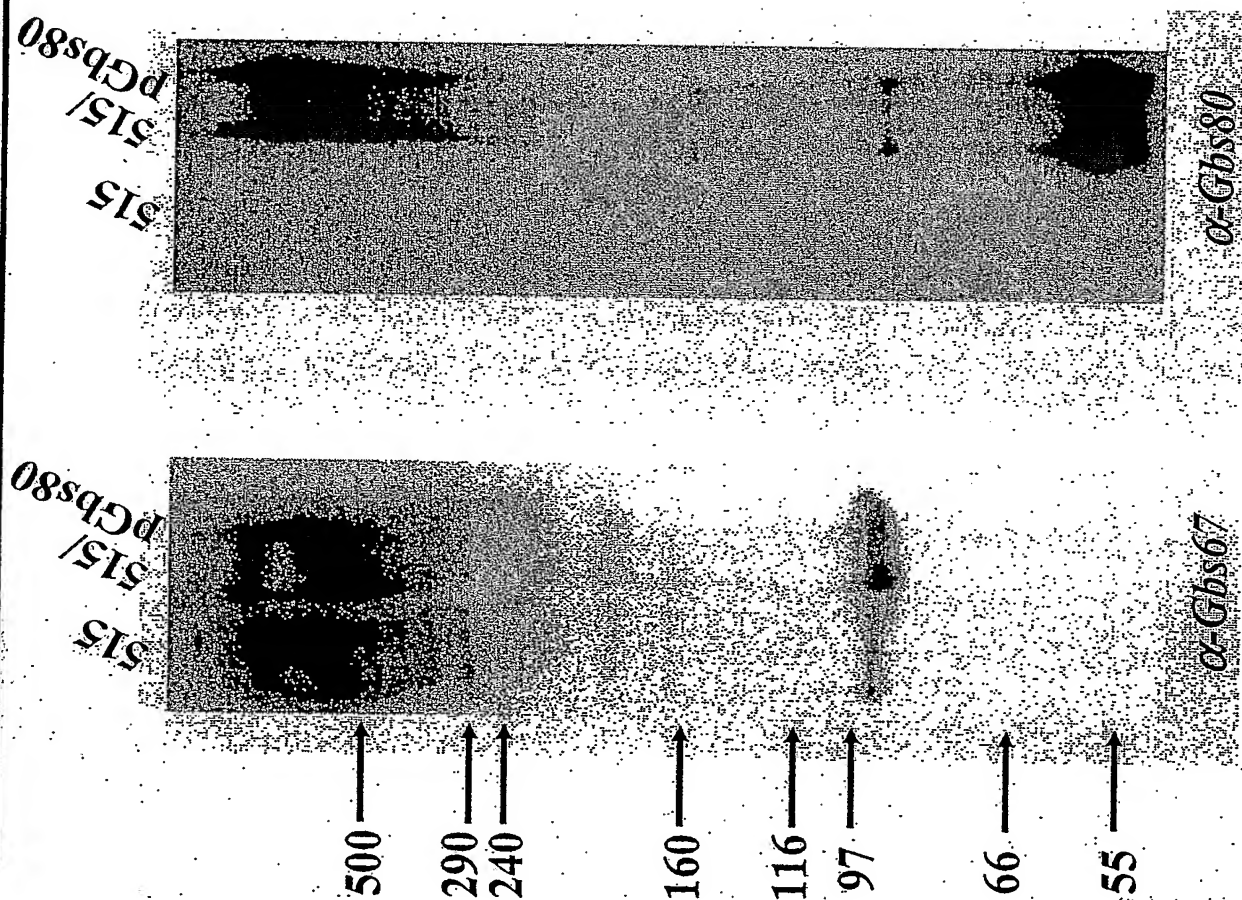
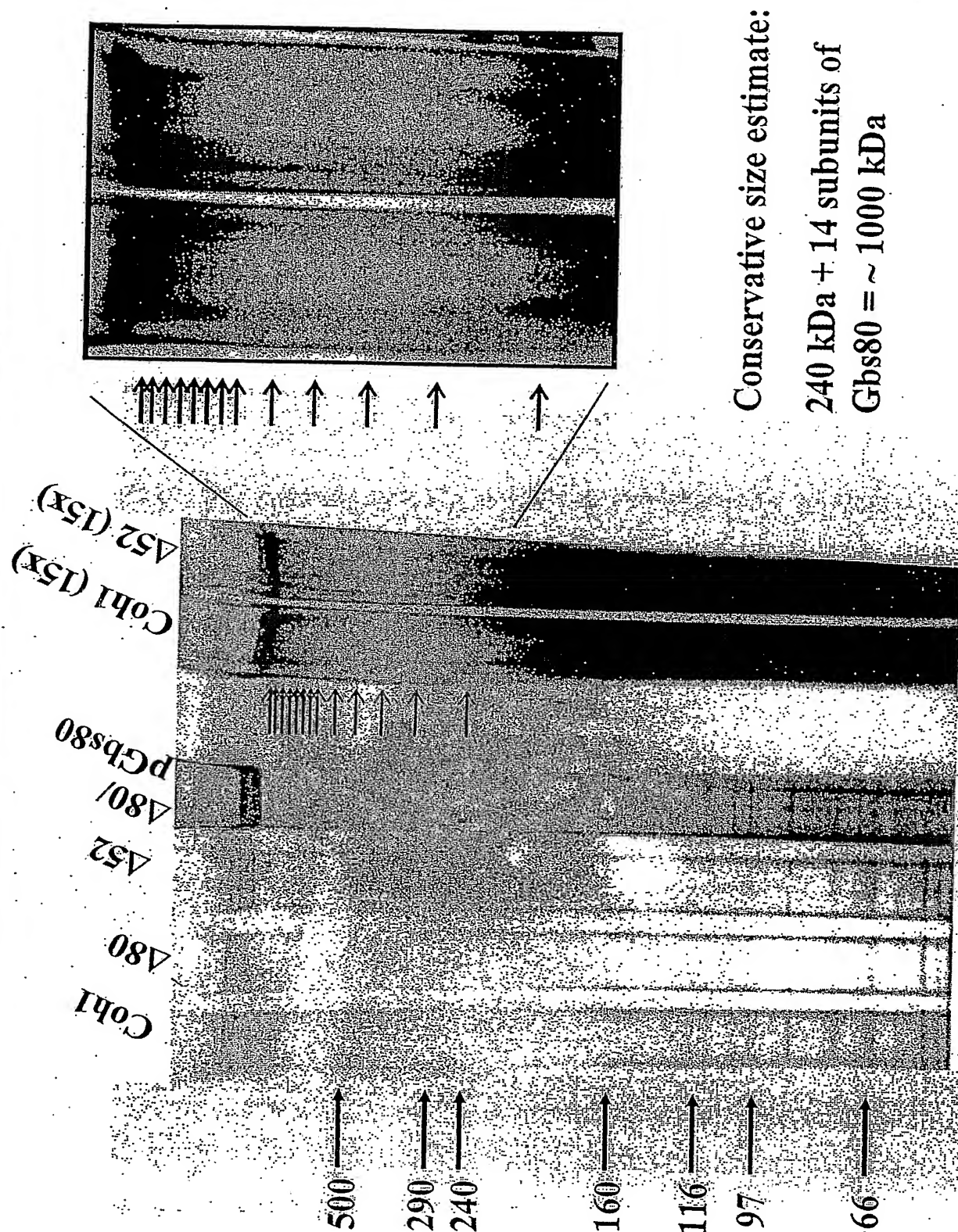
Figure 42: Gbs67 is part of a second pilus;**Gbs80 is polymerized in strain 515****(515 lacks sortase 647-8, but has AI-2 sortases)**

Figure 43: Two macro-molecules are visible in Coh1 at >1000 kDa, one is the Gbs80 pilin



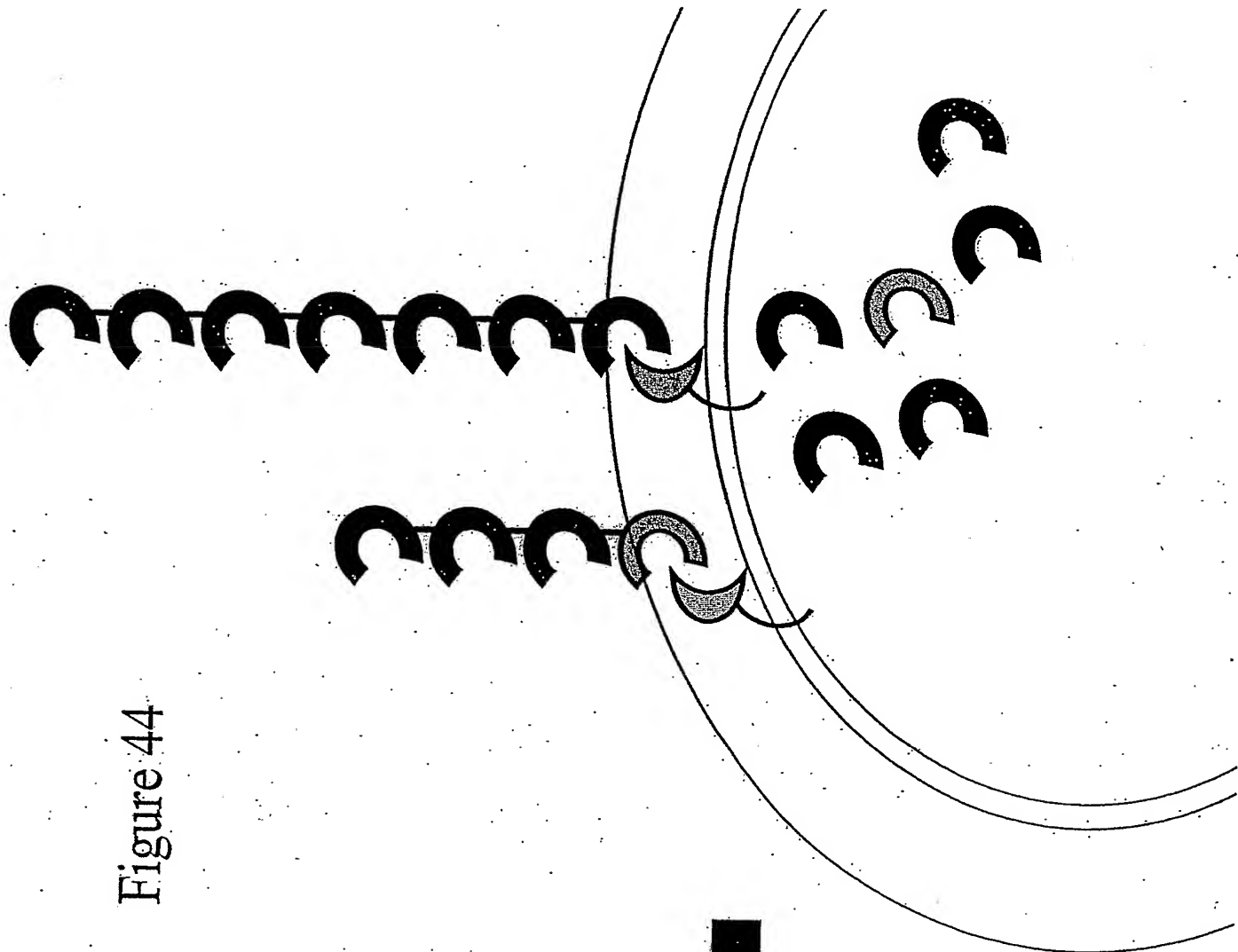


Figure 44

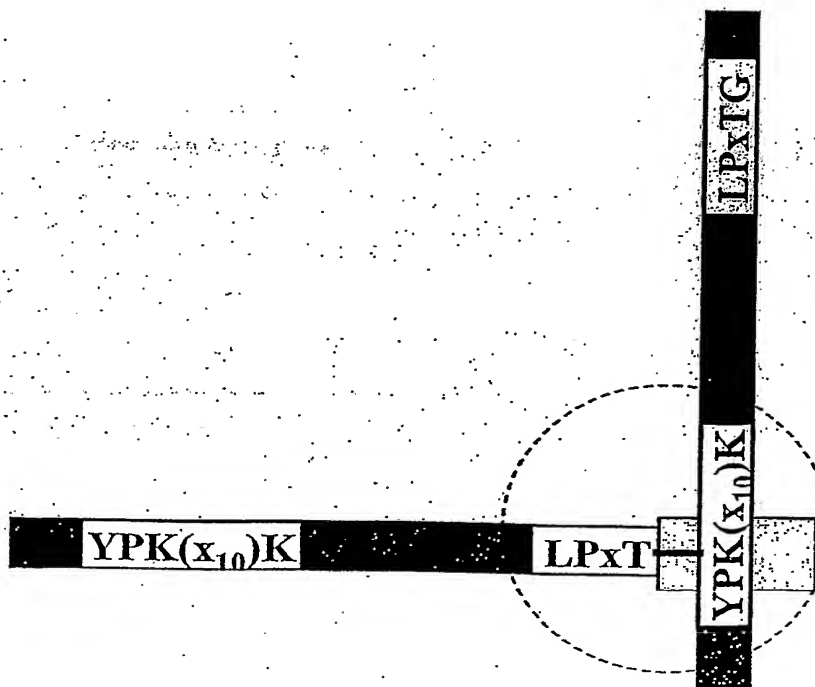


Figure 45: Gbs52 is a minor component of the GBS pilus

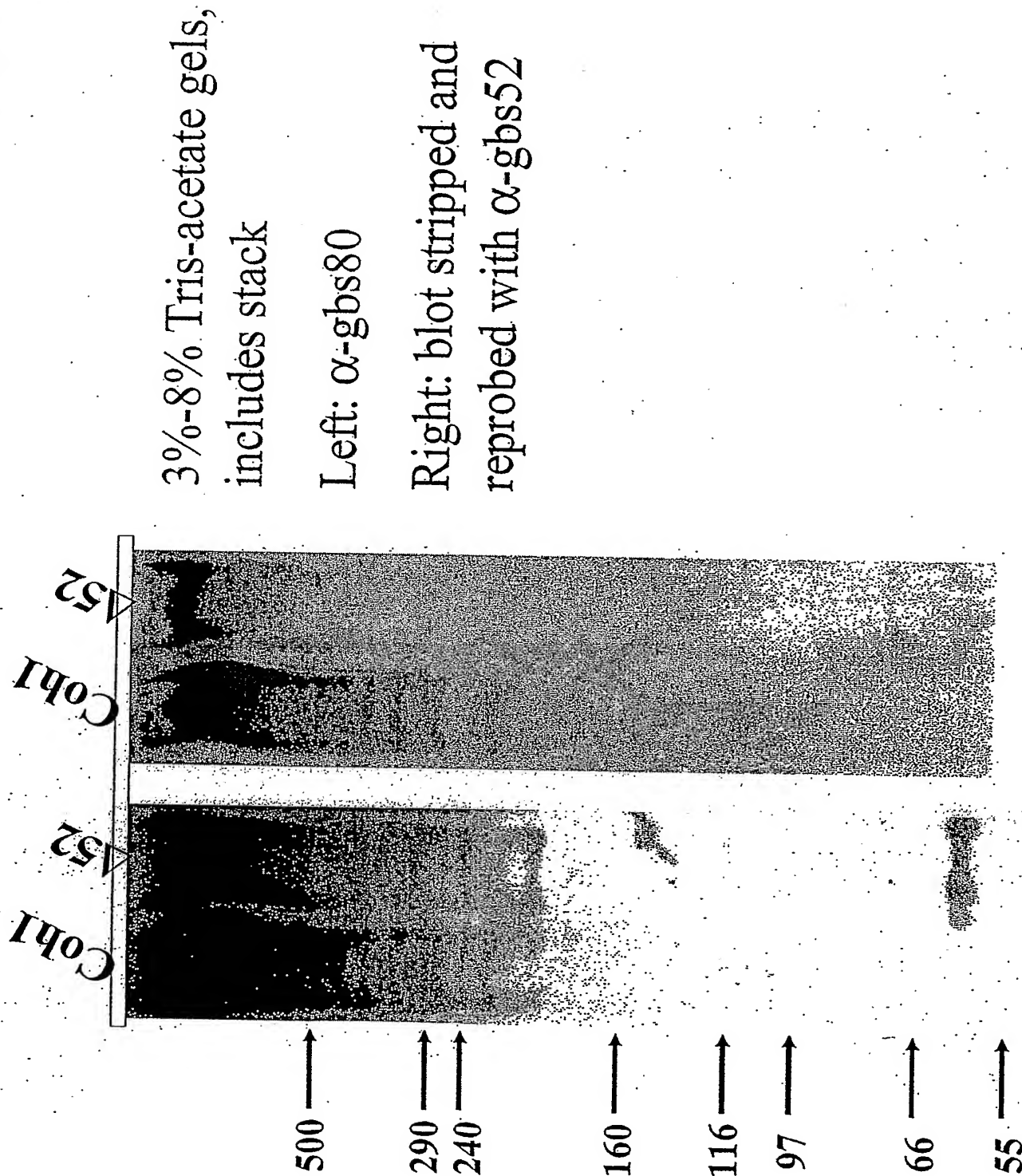


Figure 46: The pilus is found in the supernatant of the bacterial culture

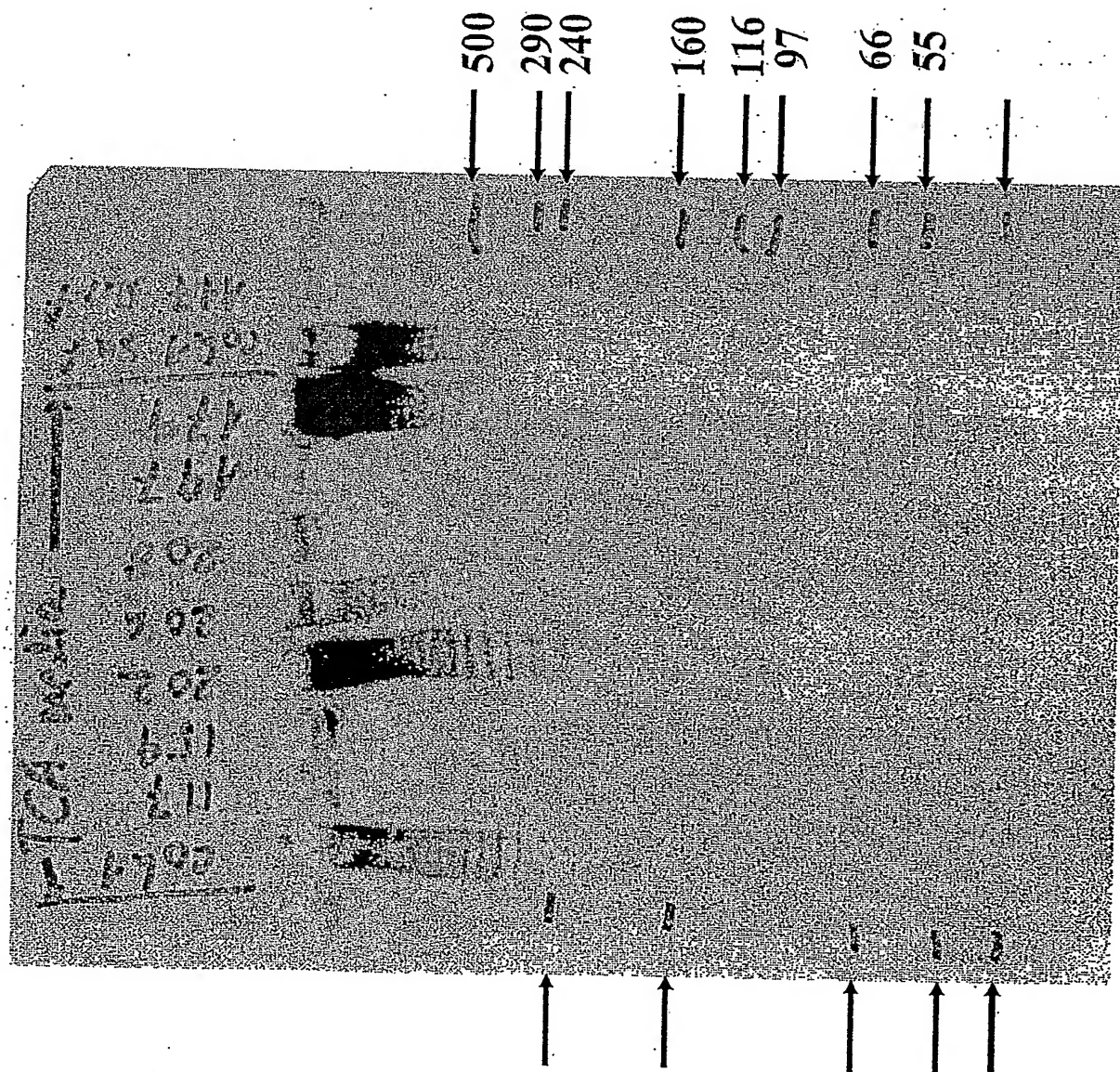
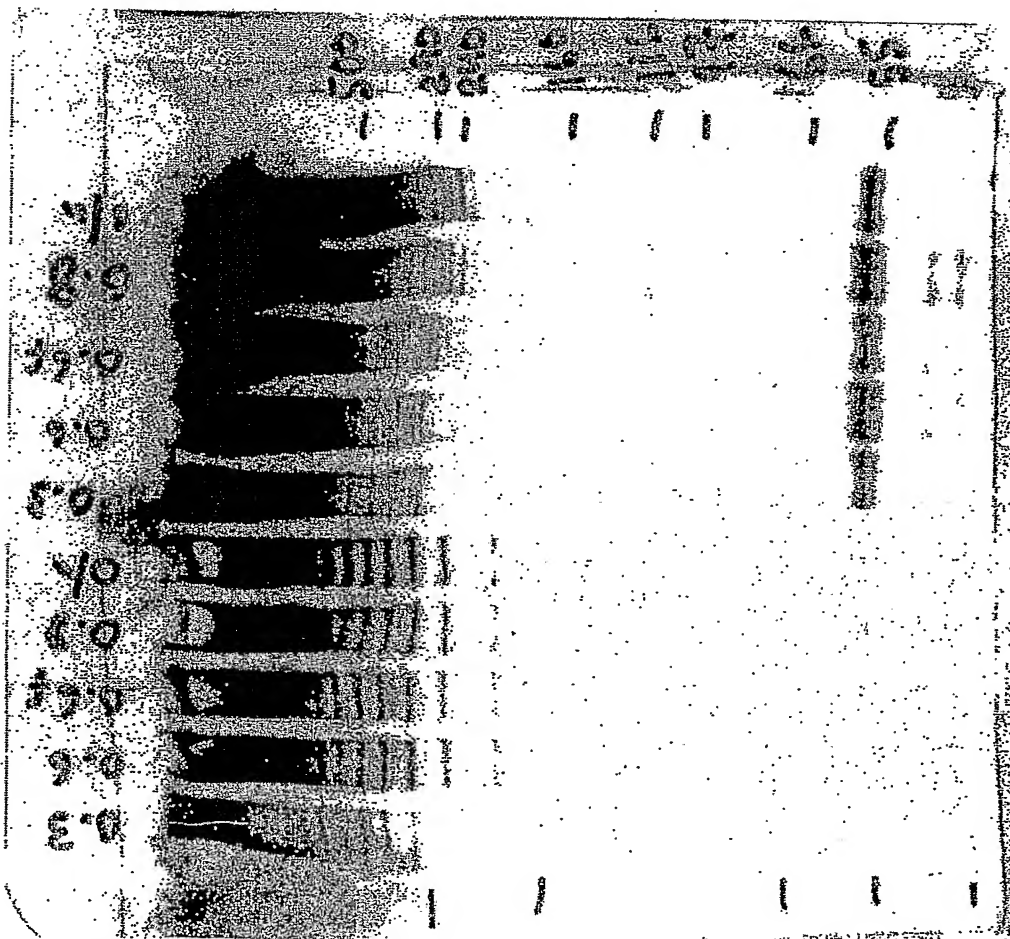


Figure 47: The pilus is found in all growth phases of cultures in all growth phases

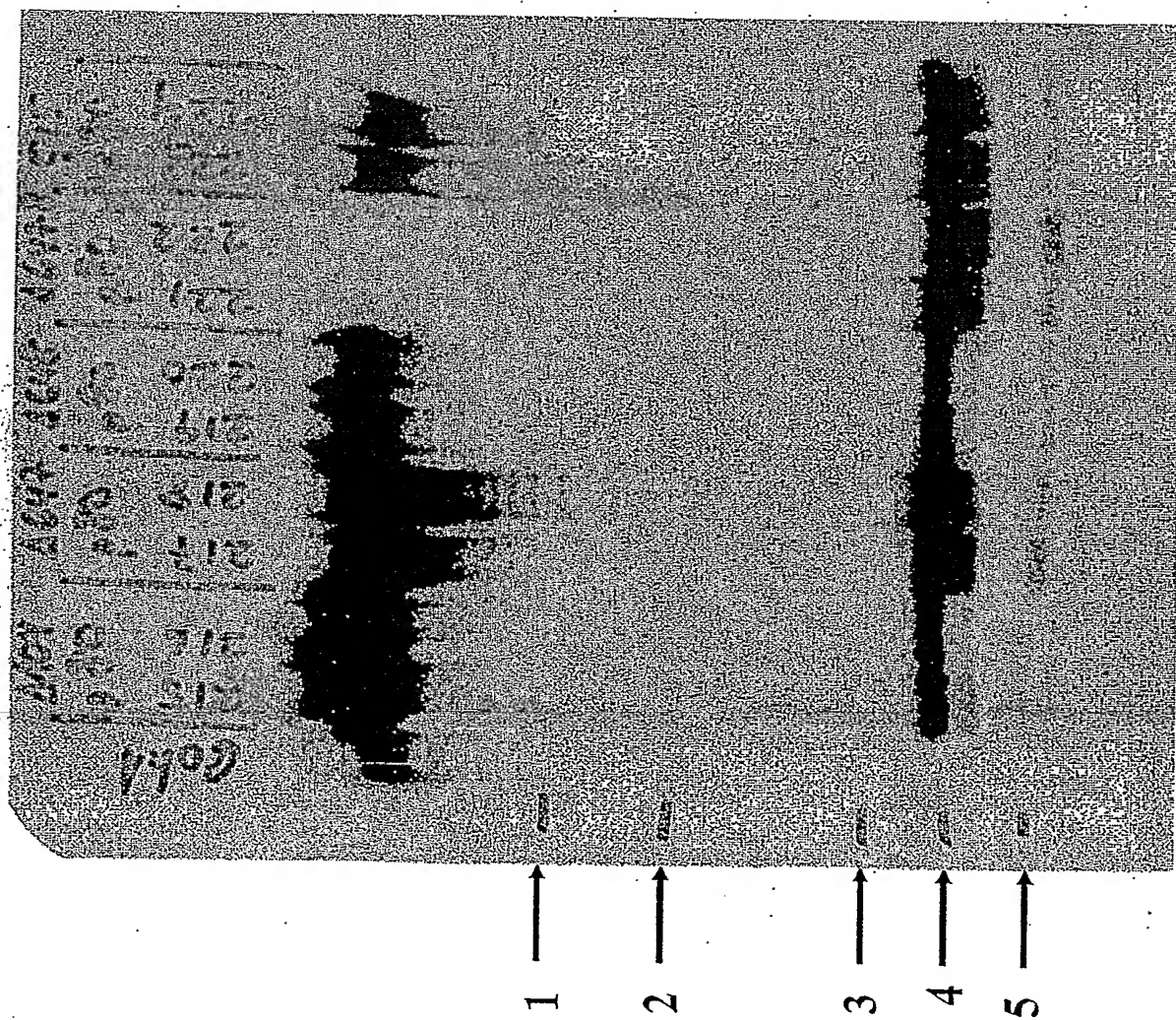


TCA precipitation of 1 ml of THB culture supernatant run on 3-8% SDS-PAGE. OD600 nm are noted above samples, "f" indicates supernatant was filtered (0.2 μ M syringe filter).

Left five samples: Coh1.

Right five samples: 179 (Δ Gbs80/pGbs80).

Figure 48: In Coh1, only the gbs80 protein and one sortase (sag0647 or sag0648) is required for polymerization



Over expression of gbs80 in various strain backgrounds (two clones each).

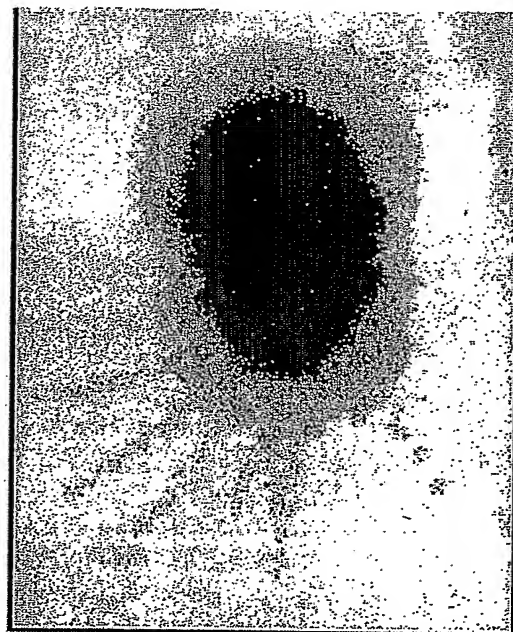
Total protein extract preparations.

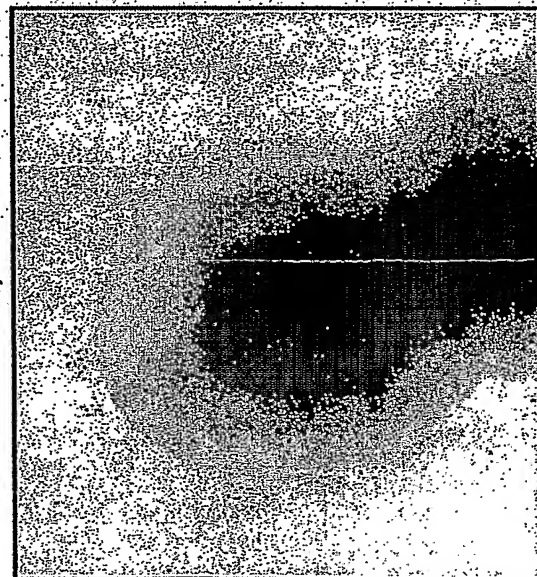
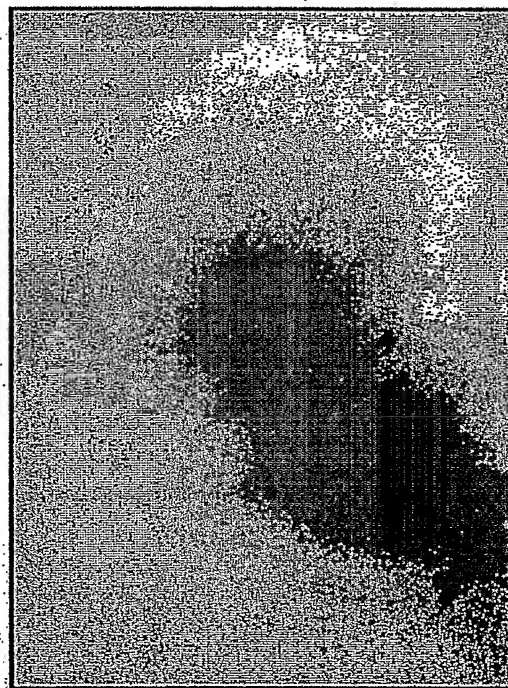
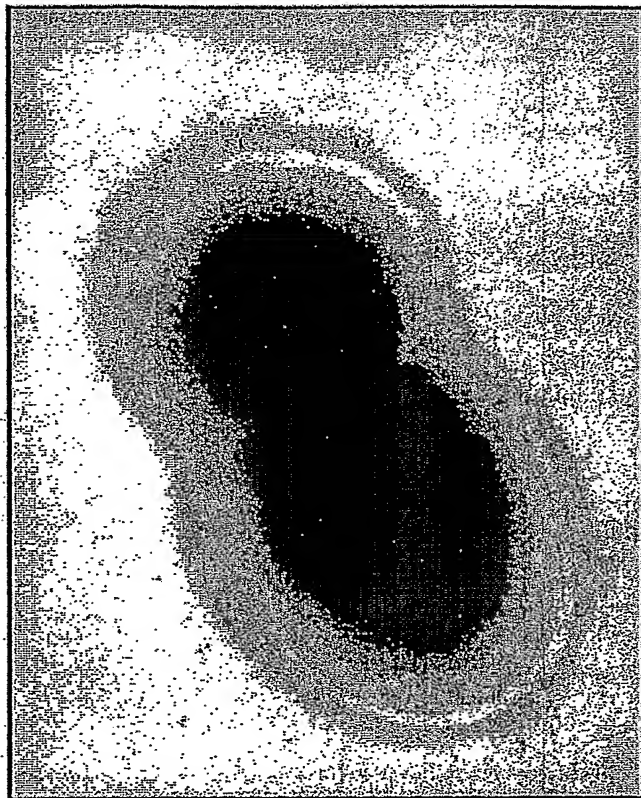
Only the double sortase mutant does not polymerize gbs80.

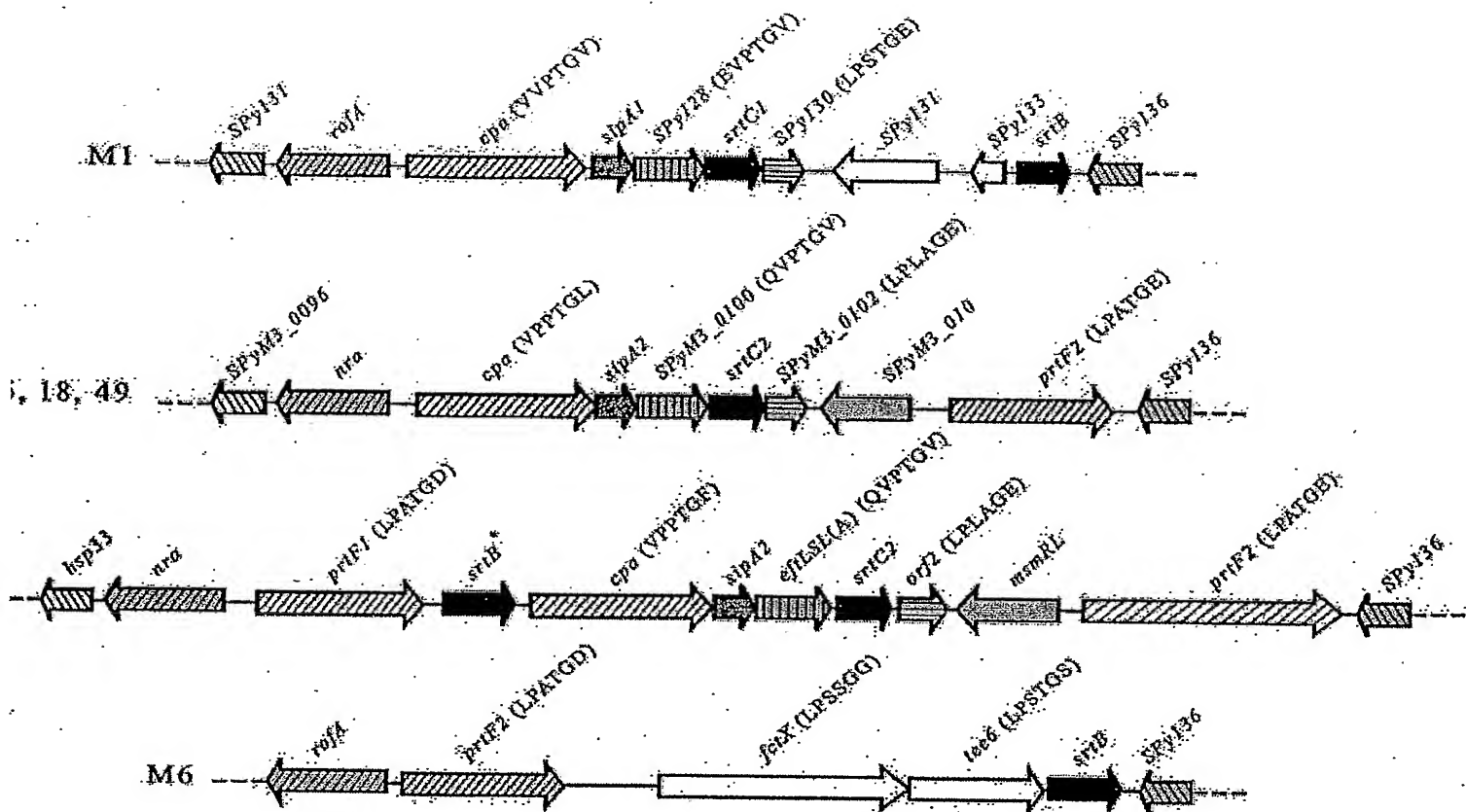
Gbs80 is polymerized in the DK515 strain background (lacks adhesin island 1, adhesin island 2 is 2603-like). Presumably, sag1405&sag1406 are responsible for polymerization.

GBS STRAIN JM9030013

IEM anti-GBS80







M1	1	MAVRUKKEDKLIIVIVIC-ETIC-LEADL-TH-
----	---	-----------------------------------

PCT/US05/27239

GI-19224135 1 MNKKKQKQDAPR-VSNRRP-----KQLTWTLGVFLMFLTLSSMRGAQSIFFGEEK
ORF78 1 -----QKRDKNYGSANNKR-----ROTTIGLLKVELTFVALIG-----IVG
GI-21909634 1 -----MOKRDKNYGSANNKR-----ROTTIGLLKVELTFVALIGIVGFSIRAFG
GI-28810257 1 -----MOKRDKNYGSANNKR-----ROTTIGLLKVELTFVALIGIVGFSIRAFG
GI-19745301 1 -----MOKRDKNYGSANNKR-----ROTTIGLLKVELTFVALIGIVGFSIRAFG
GAS15 1 LRGEKMTTRFFNKLNTLNTQRVLSKNSRRFTWTLGVFLMFLALVTSVVGAKTVFG-----

GI-19224135 53 RI[SE]VSVEKIK[SPDD]--AYPIVGYDSYDSSHPIYERFVAHDLRVNNGSRSYQVYCFNL
ORF78 39 -----
GI-21909634 46 -AEEQSVPNKQSSVQ--DYPVYGYDSYSKGYEDYSPLNTYHNKLVNLDGSREYQAYCFNL
GI-28810257 46 -AEEQSVPNKQSSVQ--DYPVYGYDSYSKGYEDYSPLNTYHNKLVNLDGSREYQAYCFNL
GI-19745301 46 -AEEQST-----
GAS15 58 -LVESTENAINPDSSESVRVYGYESYVRGHEVYKQFVAHDLRVNLEGSRSYQVYCFNL

GI-19224135 111 NSH[SE]N[SE]K[SE]F[SE]SK[SE]Q[SE]N[SE]R[SE]V[SE]D[SE]G[SE]T[SE]G[SE]V[SE]T[SE]N[SE]Y[SE]Q[SE]T[SE]P[SE]K[SE]I[SE]R[SE]G[SE]S[SE]L[SE]M[SE]K[SE]L[SE]S[SE]I[SE]M[SE]Y[SE]N[SE]A[SE]Y[SE]F[SE]K[SE]N[SE]A[SE]N[SE]G[SE]M[SE]
ORF78 39 -----
GI-21909634 103 TKHFFSKSDSVRSQWYKRLGCTNENFIKLADKPRIEDGQLOONILRIYNGYVNDNRGIM
GI-28810257 103 TKHFFSKSDSVRSQWYKRLGCTNENFIKLADKPRIEDGQLOONILRIYNGYVNDNRGIM
GI-19745301 52 -----
GAS15 117 KNAFFLGSDSSVKKYKREHDI[SE]STKEEDY[SE]M[SE]SPRI[SE]TGDELNOK[SE]TRAV[SE]M[SE]Y[SE]N[SE]G[SE]H[SE]P[SE]Q[SE]N[SE]A[SE]N[SE]G[SE]M[SE]

GI-19224135 171 DKLEPLNAILVTOQAWVYSDSSYGN-IKTLWASELKDGKIDFEQVKLMBAYSKITLSD
ORF78 39 -----ESIRAG
GI-21909634 163 KGIDPLNAILVTQNAIYVYDSSYISDTSKAFQCEETDLKIDSQOLMLRNALKRLINEK
GI-28810257 163 KGIDPLNAILVTQNAIYVYDSSYISDTSKAFQCEETDLKIDSQOLMLRNALKRLINEK
GI-19745301 52 -----
GAS15 177 EGLEPLNAIRVTQBAWVYSDNAPLISNPEDESEKRESESNLYSISQLSLNRQALKRLITEN

GI-19224135 230 LEETSKNKLFGQSKLNI[SE]V[SE]P[SE]Q[SE]D[SE]N[SE]-----VQNL[SE]LSA[SE]Y[SE]V[SE]P[SE]S[SE]P[SE]A[SE]Q[SE]S[SE]E[SE]P[SE]E[SE]V[SE]Q[SE]T[SE]K[SE]K[SE]
ORF78 46 -----AEEKSTETKKT
GI-21909634 223 EVESLPNOVPANYSLSIFQSSDNT-----FQNL[SE]LSA[SE]Y[SE]V[SE]P[SE]D[SE]T[SE]P[SE]K[SE]P[SE]G[SE]-----EPPAKTETKT
GI-28810257 223 EVESLPNOVPANYSLSIFQSSDNT-----FQNL[SE]LSA[SE]Y[SE]V[SE]P[SE]D[SE]T[SE]P[SE]K[SE]P[SE]G[SE]-----EPPAKTETKT
GI-19745301 52 -----ETKKT
GAS15 237 LATKMPKQVPDBEQLSIFQSSDNGDKYNKGQNL[SE]LSG[SE]G[SE]I[SE]V[SE]T[SE]K[SE]P[SE]E[SE]T[SE]G[SE]D[SE]P[SE]P[SE]M[SE]P[SE]N[SE]Q[SE]P[SE]O[SE]T[SE]T[SE]

GI-19224135 284 SVIIRKYAEGDYSKLEGGATLRLTGEDILFQEKVFSNGTGEKIELSNGTYTLTETSSP
ORF78 57 SVIIRKYAEGDYSKLEGGATLRLTGEDILFQEKVFSNGTGEKIELSNGTYTLTETSSP
GI-21909634 275 SVIIRKYAEGDYSKLEGGATLRLTGEDILFQEKVFSNGTGEKIELSNGTYTLTETSSP
GI-28810257 275 SVIIRKYAEGDYSKLEGGATLRLTGEDILFQEKVFSNGTGEKIELSNGTYTLTETSSP
GI-19745301 57 SVIIRKYAEGDYSKLEGGATLRLTGEDILFQEKVFSNGTGEKIELSNGTYTLTETSSP
GAS15 297 SVIIRKYAEGDYSKLEGGATLRLTGEDILFQEKVFSNGTGEKIELSNGTYTLTETSSP

GI-19224135 344 DGYKIAEPI[SE]FRV[SE]V[SE]N[SE]K[SE]V[SE]F[SE]I[SE]V[SE]Q[SE]K[SE]D[SE]G[SE]S[SE]O[SE]V[SE]E[SE]M[SE]P[SE]N[SE]K[SE]E[SE]V[SE]A[SE]P[SE

PCT/US05/27239

GI-19224135 464 ALRPRDTNPEDFLKHIKKVTEKGYNKKGD--SYNGLTEIOPRAATQLAIYYFTDSIDLRT
ORF78 233 ANTARDEDFQLFKHVNIENGYHKKGQAIPYNSLTBAQFRAATQLAIYYFTDSVDL--
GI-21909634 453 AATPRDKDADFFLKHKKILDKGYKKKGD--TYKILTEAQFRAATQLAIYYFTDSADLIT
GI-28810257 453 AATPRDKDADFFLKHKKILDKGYKKKGD--TYKILTEAQFRAATQLAIYYFTDSADLIT
GI-19745301 237 ANNPRASINDEILSQVINIVLEKGYRDDST--TYANLPSVEFRAATQLAIYYFTDSVDLON
GAS15 472 TVKPRDTDPDITFLKHKKVLEKGYREKQAIPYNSLTETQLRAATQLAIYYFTDSABL-D-

GI-19224135 522 LKTYNNGKGYHGFESWDEKTLAVTKELINYAOD--NSAPQLTNLDFFVPNNISKYQSLIGTE
ORF78 291 --TRDRLNDEHGFCDMNDOTLGVAKKIIVEYALS--DEDSKLTNLDFFVPNNISKYQSLIGTE
GI-21909634 511 LKTYNDNKGYPHGFCDKLDATLAVVHELTITYAED--VLEMTQNLDFFVPNNISKYQSLIGTO
GI-28810257 511 LKTYNDNKGYPHGFCDKLDATLAVVHELTITYAED--VLEMTQNLDFFVPNNISKYQSLIGTO
GI-19745301 295 LADY-----HGFALTTEALNATILETYAEDRANLPNISNLDFFVPNNISKYQSLIGTO
GAS15 531 ---KDKLNDYHGFCDMNDSTLAVANILMEYAO--SNPQOLTDLDFPIPNNNISKYQSLIGTO

GI-19224135 581 YHPDDLVDVIRMEDNKKQEVIPVTHSLTVKRTVVGRLGDNTNGFQFELBLNDKTSQPIVNT
ORF78 348 YHPDDLVDVIRMEDNKKQEVIPVTHSLTVKRTVVGRLGDNTNGFQFELBLNDKTSQPIVNT
GI-21909634 570 YHPNELIDVISMEDNKKQPIIPITHKLTISKTVGTIADNKKKEFNFEIHLKSSDQQAISGT
GI-28810257 570 YHPNELIDVISMEDNKKQPIIPITHKLTISKTVGTIADNKKKEFNFEIHLKSSDQQAISGT
GI-19745301 349 YHPESLVDIIRMEDNKKQPIIPITHKLTISKTVGTIADNKKKEFNFEIHLKSSDQQAISGT
GAS15 587 YHPEDLVDIIRMEDNKK--EVIPVTHSLTVKRTVVGRLGDNTNGFQFELBLNNKQELISOT

GI-19224135 641 LKTNNOQLVAKDCKYSFNLNKGDTIRIEGLEFGYSYILKETE--KDYIVTVVNNVSOEFAQS
ORF78 408 LKTNNOQLVAKDCKYSFNLNKGDTIRIEGLEFGYSYILKETE--KDYIVTVVNNVSOEFAQS
GI-21909634 630 YFTNSGELTVTDGKATFLLDGSGLIVEGLPSGYSYEITETGASDYEVSVNGKNAPDGKA
GI-28810257 630 YFTNSGELTVTDGKATFLLDGSGLIVEGLPSGYSYEITETGASDYEVSVNGKNAPDGKA
GI-19745301 409 YFTNSGELTVTDGKATFLLDGSGLIVEGLPSGYSYEITETGASDYEVSVNGKNAPDGKA
GAS15 646 VKTDKTNLEFFKDGKATINLHGESLHLOGLEFGYSYILKETE--SEGKVKVNSQEVANATV

GI-19224135 701 ASENVTADKEVTFENRKDLVPPTGLTTDCAIYLWLLLLVFFGLLWFLFGRKGLKND--
ORF78 468 ASENVTADKEVTFENRKDLVPPTGLTTDCAIYLWLLLLVFFGLLWFLFGRKGLKND--
GI-21909634 690 TKASVKEDETVAFENRKDLVPPTGLTTDCAIYLWLLLLVFFGLLWFLFGRKGLKND--
GI-28810257 690 TKASVKEDETVAFENRKDLVPPTGLTTDCAIYLWLLLLVFFGLLWFLFGRKGLKND--
GI-19745301 469 TKASVKEDETVAFENRKDLVPPTGLTTDCAIYLWLLLLVFFGLLWFLFGRKGLKND--
GAS15 706 SKTGITTSDETVAFENRKDLVPPTGLVDQKINGYLALETVAGISLCTHGTHTIRTKHD

FIGURE 52A

169/487

GI-19224134 WO 2006/078318 PCT/US2005/027239

GI-50913503 SYMFARGEKMNNTPLNKEAGFLVHTKRRKRFPAVTLVGVFLLACAGAIGFGQVAY
MVSSYMFARGEKMNNTPLNKEAGFLVHTKRRKRFPAVTLVGVFLLACAGAIGFGQVAY

GI-19224134 61 AADERTVFNFKSPDDYPPIYGYDSV-----RCIEARYHNLYVNLKGSREYQAYCFNITK
GI-50913503 61 AADERTVFNFKSPDDYPPIYGYDSV-----RCIEARYHNLYVNLKGSREYQAYCFNITK

GI-19224134 115 YEPRTYSTTNNEKKIDGSGSAFKSYAANPVLLENLDKLEKNILNVTYNGVKSNAAGF
GI-50913503 121 NNPQKNSFIKNNKKIEGCKSFVDYAHTTKEGKE---ELEQRELSELYNEYPNDANGY

GI-19224134 175 MNGLEDLNAITVTQYAVHYSDNS-QYQFETLHESEANEKISRQVTLMLREALKKLIDP
GI-50913503 178 MNGLEDLNAITVTQYAVHYSDNS-QYQFETLHESEANEKISRQVTLMLREALKKLIDP

GI-19224134 235 NLEATAANKIPSGYRLNIFKSENEAYQNLLSAEYVPPDDPPKPGDTSEHNPKTPELDGTPI
GI-50913503 237 NLEATAANKIPSGYRLNIFKSENEAYQNLLSAEYVPPDDPPKPGDTSEHNPKTPELDGTPI

GI-19224134 295 PEDPKHPDESSEPALPPLMPELDGEVPEVPSESLEPALPPLMPELDGEVPEVPSESLE
GI-50913503 297 PEDPKHPDDNLEPTLPPV-----

GI-19224134 355 PALPPLMPELDGEVPEVPSESLEPALPPLMPELDGEVPEVPSESLEPALPPLMPELDGE
GI-50913503 316 -----LDGEEVPEVPSESLEPALPPLMPELDGE

GI-19224134 415 FEVPEKPSVDLPPIEVPRYEFNNKQDQSLAGSGETEVITEVYGNQONPVDIDKKLPNETG
GI-50913503 343 FEVPEKPSVDLPPIEVPRYEFNNKQDQSLAGSGETEVITEVYGNQONPVDIDKKLPNETG

GI-19224134 475 FSGNMVETEDTKEPEVLMGGQSESVEFTKDTQTGMSCGTTPOVETEDTKEPEVLMGGQSE
GI-50913503 403 FSGNMVETEDTKEPEVLMGGQSESVEFTKDTQTGMSCGTTPOVETEDTKEPEVLMGGQSE

GI-19224134 535 SVEFTKDTQTGMSCGTTPOVETEDTKEPEVLMGGQSESVEFTKDTQTGMSCGTTPOVETE
GI-50913503 463 SVEFTKDTQTGMSCGTTPOVETEDTKEPEVLMGGQSESVEFTKDTQTGMSCGTTPOVETE

GI-19224134 595 DTKEPEVLMGGQSESVEFTKDTQTGMSCGFSETATIVEDTRPKLVFHFDDNNEPKVEENREK
GI-50913503 523 DTKEPEVLMGGQSESVEFTKDTQTGMSCGFSETATIVEDTRPKLVFHFDDNNEPKVEENREK

GI-19224134 655 PTKNITPILPATGDIENVLAFLGILILSVLSIFSLLNNKQNNKV-
GI-50913503 583 PTKNITPILPATGDIENVLAFLGILILSVLSIFSLLNNKQNNKV

FIGURE 53

170/487

GI-19745307 1 MTQKNSY FLLSLTGFILGLLLVTIGLSGVSVGHAETRNGANKQGI
ORF84 WO 2006/078318 NSYN SFLSLTGFILGLLLVTIGLSGVSVGHAETRNGANKQCAI PCT/US2005/027239
GI-28810263 1 MTQKNSYKLSFLLSLTGFILGLLLVTIGLSGVSVGHAETRNGANKQGAFEIKNN
GI-21909640 1
GI-19224141 1 MTQKNSYKLSFLLSLTGFILGLLLVTIGLSGVSVGHAETRNGANKQGSFEIKNNVDQNNKP

GI-19745307 55 ----- KSQEEYNYE
ORF84 55 ----- KSQEEYNYE
GI-28810263 55 ----- KSQEEYNYE
GI-21909640 1 -----
GI-19224141 61 LPGATFSLTSKDGKGTSVQTFSTNDKGI VDAQNLQPGTYTLKEETAPDGYD KSRITVT

GI-19745307 64 VYDN ----- RNI
ORF84 64 VYDN ----- RNI
GI-28810263 64 VYDN ----- RNI
GI-21909640 1 -----
GI-19224141 121 VYENG YTKLVENPYNGEII SKAGSKDVSSSLQLENPKMSVVS KYGKTEVSSGAADFY RNI

GI-19745307 71 LODGEHKLEIKRVDGTGKTYQG FCFQLTKNFF TAQGVSKKLYKKLSS
ORF84 71 LODGEHKLEIKRVDGTGKTYQG FCFQLTKNFF TAQGVSKKLYKKLSS
GI-28810263 71 LODGEHKLEIKRVDGTGKTYQG FCFQLTKNFF TAQGVSKKLYKKLSS
GI-21909640 1 MSS
GI-19224141 181 AAYFKMSF EKKOK E KSETIN E GDT E V L Q L D R L N P K G I S Q D E P I T Y D S A N S P L A I G K Y H

GI-19745307 118 ----- SDEETLK
ORF84 118 ----- SDEETLK
GI-28810263 118 ----- SDEETLK
GI-21909640 4 ----- SDEETLK
GI-19224141 241 AENHQLIYTFTDYIAGLDKVLQSLAELSLFLENK E V L E N T S I S N F K S T I G G Q E I T Y K G T V N

GI-19745307 125 QYASKYTSNRRGDTSC
ORF84 125 QYASKYTSNRRGDTSC
GI-28810263 125 QYASKYTSNRRGDTSC
GI-21909640 11 QYASKYTSNRRGDTSC
GI-19224141 301 VLYGNESTKESNYHTNGLSNVCGSIESYNTETGEFVWVYVNPNTNI PYATMNLWGFR

GI-19745307 141 ----- NLKKQIAKVLTEGYPT
ORF84 141 ----- NLKKQIAKVLTEGYPT
GI-28810263 141 ----- NLKKQIAKVLTEGYPT
GI-21909640 27 ----- NLKKQIAKVLTEGYPT
GI-19224141 361 ARSNTSDLENDANTSSAELGEIOVYEVPEGEKLPSSYGVDVTKL L R T L I T A G L N G F O M

GI-19745307 157 NKS DWLN GLTENENIEVTQDAIWF
ORF84 157 NKS DWLN GLTENENIEVTQDAIWF
GI-28810263 157 NKS DWLN GLTENENIEVTQDAIWF
GI-21909640 43 NKS DWLN GLTENENIEVTQDAIWF
GI-19224141 421 TTRORI DFCNNIQNKAFI I KVTGKTQSGKPLVVQSNLASFRGASEYAAFTPVGGN VYEQ

GI-19745307 182 TETTVPADR SYTNRNVNVSQRMKEVYQKLIDTTDID KYEDVQFDLFVPQDTN
ORF84 182 TETTVPADR SYTNRNVNVSQRMKEVYQKLIDTTDID KYEDVQFDLFVPQDTN
GI-28810263 182 TETTVPADR SYTNRNVNVSQRMKEVYQKLIDTTDID KYEDVQFDLFVPQDTN
GI-21909640 68 TETTVPADR SYTNRNVNVSQRMKEVYQKLIDTTDID KYEDVQFDLFVPQDTN
GI-19224141 481 NEIALSPSKGSGSGKSPETKPSITVANLKRVAQLRFKKMSTDNVPLPEAAFE LRSSNGNS

GI-19745307 233 LQAVISVEPVIESLPITS LKPIAQNDITAKK
ORF84 233 LQAVISVEPVIESLPITS LKPIAQNDITAKK
GI-28810263 233 LQAVISVEPVIESLPITS LKPIAQNDITAKK
GI-21909640 119 LQAVISVEPVIESLPITS LKPIAQNDITAKK
GI-19224141 541 QKLEASSNTQGEVHFKDLTS GTYDLYETKAPKGYQQVTEKLEATVTVDTTNPAEEMVTWGS

FIGURE 54